# Unbounded-rate Markov decision processes:
## structural properties via a parametrisation approach

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 23 juni 2016
klokke 15:00 uur

door

Herman Blok

geboren te Utrecht in 1985

Samenstelling van de promotiecommissie:

**Promotor:**
Prof. dr. W.Th.F. den Hollander

**Copromotor:**
Dr. F.M. Spieksma

**Overige Leden:**
Prof. dr. J.J. Meulman (Universiteit Leiden, secretary)
Prof. dr. ir. I.J.B.F. Adan (Technische Universiteit Eindhoven)
Prof. dr. N. Bäuerle (Karlsruhe Institute of Technology)
Dr. S. Bhulai (Vrije Universiteit Amsterdam)

Omslagillustratie: Kaartverkoop Olympische Spelen 1928 in Amsterdam. Duizenden mensen staan uren in de rij voor een toegangsbewijs voor de voetbalwedstrijd Nederland-Uruguay. Vijzelstraat Amsterdam, Nederland. (Nationaal Archief/Collectie Spaarnestad/Fotograaf onbekend)

# Contents

# 1 Introduction

## 1.1 Motivation

Central in this thesis are countable state continuous time Markov decision processes. Most problems treated here are related to the domain of queueing theory. Queueing problems have been the main motivation for the conducted research, in particular problems with unbounded transition rates. These rates may be unbounded due to an infinite server pool (see Chapter 4), due to reneging customers (see Chapter 5) or due to other reasons. Naturally, Markov decision theory has a wider range of applications to which the results apply as well.

Our interest is in structural properties of optimal policies in Markov decision processes. In finite state problems the boundary can potentially break the monotonicity of the structural properties. Therefore we focus on problems with an infinitely countable state space, since optimal policies for these problems exhibit better structural properties. Countable state problems can be seen as a good approximation of problems with a large state space. Knowledge about the structure of an optimal policy reduces the class of potentially optimal policies.

There are several ways to derive such structures. One of the most powerful methods is to use monotonicity properties of the value function and then show the structure of an optimal policy via the optimality equation. This thesis is concerned with the question of how to show such monotonicity properties. For discrete time Markov decision processes the literature provides a satisfying answer to this question, in the form of value iteration. However, this is not the case for continuous time processes. Especially if a Markov decision process has unbounded jump rates – so it does not allow uniformisation – there are serious shortfalls in the theory. We refer to Section 2.1 for an extensive discussion of the difficulties and the gaps in the literature.

Chapters 2 and 3 deal with some of these gaps, building towards a systematic approach to derive structural properties. Chapters 4 and 5 use the developed methods to analyse queueing models with unbounded transition rates. In Chapter 6 we illustrate how different truncations can have an effect on the preservation of structural properties. In the framework of event based dynamic programming, Chapter 7 provides a systematic overview of propagation results that can be used for the analysis of rate truncated models.

## 1.2 Structure of the thesis

Below we provide a more detailed overview of this thesis.

Chapter 2 gives a systematic overview of the derivation of structural properties for both discrete and continuous time, and for both the $\alpha$-discounted and the average cost criteria. The emphasis in this chapter is to provide roadmaps for such derivations. Essential to these roadmaps is the reduction to the 'base case' model: discounted cost optimality in discrete time. We study conditions justifying the vanishing discount approach for continuous time, a method to obtain results for the average cost criterion by letting the discount factor vanish. Together with the results derived in Chapter 3 this makes the reduction to the base case possible. This chapter is based on Blok and Spieksma [19].

Chapter 3 treats parametrised Markov processes in continuous time with a discounted cost. Under mild drift conditions, we derive continuity of the value function as a function of the parameter. The parameter may represent a perturbation of a Markov decision process, thus allowing that a Markov decision process with unbounded rates may be a approximated by a sequence of uniformisable Markov decision processes. The framework of parametrised Markov processes is illustrated by an application to the discounted cost version of the server farm model of [1]. Chapter 3 is based on Blok and Spieksma [18].

In Chapter 4 we analyse the server farm model in more detail. The goal is to reduce the cost for servers that are unnecessarily idle, while keeping the penalty cost for not having any directly available servers upon customer arrival low as well. The framework developed in Chapters 2 and 3 allows to extend the results for the bounded rate server farm model of Adan et al. [1]. This yields optimality of a switching curve under both the $\alpha$-discounted and average cost criteria. The special structure of this problem allows to reduce the problem to solving a nested sequence of finite rate and action space problems. Using coupling methods additional properties of the model can be derived including existence of a Blackwell optimal policy. This chapter is based on Blok et al. [17].

Chapter 5 treats the $K$-competing queueing problem with customer abandonment. To analyse the model we use the truncation principle called *smoothed rate truncation* invented by Bhulai et al. [11]. This advanced truncation technique has the ability to preserve structural properties after truncation. The result of the analysis is optimality of a simple index policy that can be seen as a generalisation of the classical $c\mu$-rule. This chapter is based on Bhulai et al. [10].

In Chapter 6 we discuss truncation techniques of a service allocation problem in a tandem queue. A priority rule is proven to be optimal for certain parameter conditions. However, a numerical calculation using a straightforward truncation exposes a strikingly different policy. We show how other truncations can preserve the optimal policy.

Chapter 7 provides a systematic list of propagation results in the framework of event based dynamic programming. The propagation results form the building blocks for propagation through value iteration. It requires the introduction of operators describing the effect of different events related to queues with unbounded rates and to rate truncations. Most of the operators and propagations have already appeared in Chapters 4 and 5 or in [16].

# 2 A roadmap to structures for Markov decision processes

This chapter is based on Blok and Spieksma [19], submitted.

## 2.1 Introduction

The question how to rigorously prove structural results for continuous time Markov decision problems (MDPs) with a countable state space and unbounded jump rates (as a function of state) seems to be an assiduous task. As a typical example one may consider the competing queues model with queue dependent cost rates per customer and per unit time, where the objective is to determine the server allocation that minimises the total expected discounted cost or expected average cost per unit time. Both discounted and average cost are known to be minimised by the $c\mu$-rule, which prescribes to allocate the server to the queue that yields the largest cost reduction per unit time. A possible method to tackle this problem is to apply value iteration (VI) to the uniformised discrete time MDP and show that optimality of the $c\mu$-rule propagates through the VI step.

If customer abandonment is allowed, the resulting MDP is a continuous time MDP with unbounded jumps, since customers may renege at a rate that is proportional to the number of customers present in each of the queues. In order to apply VI, one needs to time-discretise the MDP. One way to associate a discrete time MDP with this problem is by constructing the decision process embedded on the jumps of the continuous time MDP. However, it is not clear whether structural properties propagate through the VI step (cf. Section 2.3.4). Another solution is to perturb or truncate the continuous time MDP, so that it becomes uniformisable and then apply VI. A suitable truncation or perturbation needs to be invariant with respect to structural properties of interest of the investigated MDP.

The first question is whether there exists generic truncation methods that possess such an invariance property. Clearly, this can never be systematically proved, since it depends on the properties that one wishes to prove. However, one might be able to formulate recommendations as to what kind of perturbation methods perform well, with regard to such an invariance requirement.

The paper [27] studies two competing queues with abandonments, and a problem-specific truncation is used. Later [11] has introduced a truncation method, called smoothed rate truncation (SRT) that so far seems to work well for problems where a simple bounded rate truncation (as in Section 2.3.4) does not work. In addition, it can be used for numerical applications in bounded rate optimisation problems (cf. Section 2.2.2). The SRT method has been used in Chapter 5 for identifying conditions under which a simple index policy is optimal in the competing queues problem with abandonments.

Consecutively, suppose that an application of the truncation method yields truncated MDPs with optimal policies and a value function that have the properties of interest. For instance, the above mentioned application of SRT to the competing queues example with abandoning customers yields optimality of an index policy for each truncated MDP. However, these properties still have to be shown to apply to the original non-truncated problem. Thus, convergence results are required that yield continuity of the optimal policy and value function in the truncation parameter, in order to deduce the desired results for the non-truncated MDP from the truncated ones.

A second question therefore is as to what kind of easily verifiable conditions on the input parameters of the perturbed MDP guarantee convergence of the value function and optimal policy to the corresponding ones of the original unperturbed MDP. In [27], the authors had to prove a separate convergence result apart from devising a suitable truncation and prove that VI propagates the properties of interest. Apparently, theory based on a set of generic conditions that can incorporate convergence within the optimisation framework was lacking. This lack is precisely what has hampered the analysis of the server farm model in [1], where the authors have restricted their analysis to showing threshold optimality of a bounded rate perturbed variant of the original model. Apparently no appropriate tools for deducing threshold optimality of the original unbounded problem from the results for the perturbed one were available to them.

A third major problem occurs in the context of the average cost criterion. In particular, VI is not always guaranteed to converge. This is true under weak assumptions in discounted cost problems, however, in average cost problems there are only limited convergence results (cf. Section 2.2.3). One of these requires strong drift conditions, that do not allow transience under any stationary deterministic policy. However, often this is not a convenient requirement. One may get around this difficulty by a vanishing discount approach, which analyses the expected average cost as a limit of expected $\alpha$-discounted costs as the discount factor tends to 0 (or 1, depending on how the discount factor is modelled).

For a model like the competing queues model with abandonments, a multistep procedure to obtain structural results for the average cost problem then would be as follows. First, consider the $\alpha$-discounted cost truncated problem. Structural results for the $\alpha$-discounted cost non-truncated problem follow, by taking the limit for the truncation parameter to infinity. Finally, taking the limit of the discount factor to 0 hopefully yields the final structural results for the original continuous time average cost problem.

For some of these steps theoretical validation has been provided for in the literature, but not for all, and not always under conditions that are easily checked. The main focus of this chapter is to fill some gaps in the described procedure, whilst requiring conditions that are formulated in terms of the input parameters. Based on the obtained results, we aim to provide a systematic and feasible approach for attacking the validation of structural properties, in the spirit of the multistep procedure sketched above. We hope that this multistep procedure will also be beneficial to other researchers as a roadmap for tackling the problem of deriving structural results for problems modelled as MDPs.

We do not address the methods of propagating structures of optimal policies and value function through the VI step. Such methods belong to the domain of 'event based dynamic programming', and they have been discussed thoroughly in [45], with extensions to SRT and other rate truncations in [16] and Chapter 7. Furthermore, we do not include an elaborate evaluation of closely related results from the literature. Some detailed comments has been included in this chapter, whenever we thought it relevant.

Another omission in this work is the study of perturbed MDPs with the average cost criterion. However, the conditions required for achieving the desired continuity results as a function of a perturbation parameter are quite strong. Therefore a more recommendable approach would be the one we have developed in this chapter, using the vanishing discount approach. As a last remark: we generally restrict to the class of stationary policies, and not history-dependent ones. Especially the results quoted for discrete time MDPs apply to the larger policy class. In continuous time MDPs allowing history-dependent policies causes extra technical complications that we do not want to address in this work.

A short overview of the chapter content is provided next. In Section 2.2 we discuss discrete time, countable state MDPs, with compact action sets. First, the $\alpha$-discount optimality criterion is discussed, cf. Section 2.2.1. This will be the base case model, to which the MDP problems might have to be reduced in order to investigate its structural properties. We therefore describe it quite elaborately. In addition, we have put it into a framework that incorporates truncations or perturbations. We call this a parametrised Markov process. Interestingly enough, 'standard' but quite weak drift conditions introduced for $\alpha$-discounted cost MDPs in discrete time, allowed this extension to parametrised Markov processes, with no extra effort and restriction. It incorporates the finite state space case, elaborated on in the seminal book [26].

In Section 2.2.2 we provide a discussion of SRT, as a method for numerical investigation of structural properties of a countable state MDP. The conditions that we use are a weak drift condition on the parametrised process, plus reasonable continuity conditions. This has been based on the work in [47, 75] for MDPs.

In Section 2.2.3 we study the expected average cost criterion, whilst restricting to non-negative cost, i.e. negative dynamic programming. This restriction allows transience, and the analysis follows [21], in the form presented by [63]. Basically, the conditions imposed require the existence of one 'well-behaved' policy, and a variant of inf-compact costs. The latter ensures that optimal policies have a guaranteed drift towards a finite set of low cost states. The contribution of these works is that they validate the vanishing discount approach, thus allowing to analyse the discrete time average cost problem via the discrete time $\alpha$-discounted cost problem.

Then we turn to studying continuous time MDPs in Section 2.3. First the $\alpha$-discounted cost problem is considered. The drift conditions on parametrised discrete time Markov processes have a natural extension to continuous time. The results listed are based on Chapter 3, but the literature contains quite some work in the same spirit within the framework of MDPs with more general state spaces, cf. e.g. [35, 53], and references therein. A closely related perturbation approach has been studied in [52]. Since perturbations are incorporated in the parametrised framework, the approach allows to study bounded jump perturbations. Indeed, optimal policies and value functions are continuous as a function of the perturbation parameter. In this way, Chapter 3 obtains threshold optimality of the original unbounded $\alpha$-discounted cost variant of the server farm model studied in [1].

Finally, for the expected average cost criterion, we use the natural generalisation of the discrete time conditions. Although closely related to analyses in [35, 53] and references therein, as far as we know this form has not appeared yet in the literature. The vanishing discount approach is validated in the same way as was done for the discrete time MDP. This reduces the problem of studying structural properties for average cost MDPs, satisfying the proposed conditions, to analysing a continuous time $\alpha$-discounted cost MDP, for which the solution method has already been described. As a consequence, also average cost threshold optimality for the above mentioned server farm model from [1] follows from $\alpha$-disount optimality of a threshold policy, cf. Chapter 4.

Dispersed through the chapter are roadmaps for attacking the validation of structural properties. These are summarised in Section 2.3.4.

## 2.2 Discrete time Model

In this section we will set up a framework of parametrised Markov processes in discrete time. With an extra assumption – the product property – a parametrised Markov process reduces to a discrete time MDP. However, treating this in the parametrised framework allows for results on perturbations or approximations of MDPs as well. Notice that instead of the usual nomenclature 'Markov chain' for a Markov process in discrete time, we will consistently use 'Markov process', whether it be a process in discrete or continuous time.

Let $\Phi$ be a parameter space. Let $\mathbf{S}$ denote a countable space. Each parameter $\phi \in Phi$ is mapped to an $\mathbf{S}\times\mathbf{S}$ stochastic matrix $P(\phi)$, and a cost vector $c(\phi) : \mathbf{S} \to \mathbb{R}$. We denote the corresponding elements by $p_{xy}(\phi)$, $x, y \in \mathbf{S}$ and $c_x(\phi)$, $x \in \mathbf{S}$. If $f : \mathbf{S} \to \mathbb{R}$, then $P(\phi)f$ is the function with value

$$P(\phi)f_x = \sum_y p_{xy}(\phi)f_y$$

at point $x \in \mathbf{S}$, provided the integral is well-defined.

To transition matrix $P(\phi)$ one can associate a Markov process on the path space $\Omega = \mathbf{S}^\infty$. Given an initial distribution $\nu$ on $\mathbf{S}$, the Kolmogorov consistency theorem (see e.g. [14]) provides the existence of a probability measure $\mathsf{P}_\nu^\phi$ on $\Omega$, such that the canonical process $\{X_n\}_n$ on $\Omega$, defined by

$$X_n(\omega) = \omega_n$$

is a Markov process with transition matrix $P(\phi)$, and probability distribution $\mathsf{P}_\nu^\phi$. The corresponding expectation operator is denoted by $\mathsf{E}_\nu^\phi$. To avoid overburdened notation, we have put the dependence on the parameter $\phi$ in the probability and expectation operators, and not in the notation for the Markov process. We further denote $P^{(n)}(\phi)$ for the $n$-th iterate of $P(\phi)$, where $P^{(0)}(\phi) = \mathbf{I}$ equals the $\mathbf{S} \times \mathbf{S}$ identity matrix.

We assume the following basic assumption.

**Assumption 2.2.1.** The following conditions hold:

**i)** the parameter space $\Phi$ is locally compact;

**ii)** $\phi \mapsto p_{xy}(\phi)$ continuous on $\Phi$ for each $x, y \in \mathbf{S}$;

**iii)** $\phi \mapsto c_x(\phi)$ is continuous on $\Phi$ for each $x \in \mathbf{S}$.

To incorporate MDPs in this set up, we use the following concept.

**Definition 2.2.1.** Let $\Phi' \subset \Phi$, inheriting the topology on $\Phi$. We say that $\{P(\phi), c(\phi)\}_{\phi \in \Phi'}$ has the product property with respect to $\Phi'$ if

**i)** there exist compact sets $\Phi'_x$, $x \in \mathbf{S}$, such that $\Phi' = \prod_{x \in \mathbf{S}} \Phi_x$; then $\Phi'$ is compact in the product topology;

**ii)** for any $\phi, \phi' \in \Phi'$, $x \in \mathbf{S}$ with $\phi_x = \phi'_x$, it holds that

- $(\,P(\phi))_{x\cdot} = (\,P(\phi'))_{x\cdot}$, where $(\,P(\phi))_{x\cdot}$ stands for the $x$-row of $P(\phi)$;
- $c_x(\phi) = c_x(\phi')$.

For notational convenience we will simply say that $\Phi'$ has the product property. Under the product property, with a slight abuse of notation we may write $c_x(\phi_x)$ and $p_{xy}(\phi_x)$ instead of $c_x(\phi)$ and $p_{xy}(\phi)$. In case the dependence on $\phi$ is expressed in the probability or expectation operators, we write $c_{X_n}$ instead $c_{X_n}(\phi)$.

*Remark* 2.2.1. If $\Phi$ has the product property, then the parametrised Markov process is an MDP. The set $\Phi$ may represent the collection of deterministic stationary policies, and we will denote it by $\mathcal{D}$. In this case $\mathcal{D}_x$ is the action set in state $x \in \mathbf{S}$.

For any $x \in \mathbf{S}$, let $\pi_x$ by a probability distribution on $\mathcal{D}_x$. Then $\pi = (\pi_x)_x$ is a stationary, randomised policy. The collection $\Pi$ of all stationary randomised policies can be viewed as a parameter set having the product property as well. We will not consider this explicitly, but all discussed results cover this case as well.

Next we define the various performance measures and optimality criteria that we will study. Lateron we will provide conditions under which these are well-defined, and optimal polices exist.

For $0 < \alpha < 1$, define the *expected total $\alpha$-discounted cost value function* $v^\alpha(\phi)$ under parameter $\phi \in \Phi$ by

$$v_x^\alpha(\phi) = \mathsf{E}_x^\phi\Big[\sum_{n=0}^\infty (1-\alpha)^n c_{X_n}\Big], \quad x \in \mathbf{S}. \tag{2.1}$$

Notice that the discount factor is taken to be $1-\alpha$. Usually the discount factor is taken to equal $\alpha$ instead. Our choice here allows a more direct analogy with the continuous time case.

Next let $\Phi' \subset \Phi$ have the product property. Define the *minimum expected total $\alpha$-discounted cost* $v^\alpha$ w.r.t. $\Phi'$ by

$$v^\alpha_x = \inf_{\phi\in\Phi'} \{v^\alpha_x(\phi)\}, \quad x \in \mathbf{S}.$$

If for some $\phi \in \Phi'$ it holds that $v^\alpha = v^\alpha(\phi)$, then $\phi$ is said to be $\alpha$-discount optimal (in $\Phi'$).

The *expected average cost* $\mathsf{g}(\phi)$ under parameter $\phi \in \Phi$ is given by

$$\mathsf{g}_x(\phi) = \limsup_{N\to\infty} \frac{1}{N+1} \mathsf{E}^\phi_x\Big[\sum_{n=0}^{N} c_{X_n}\Big], \quad x \in \mathbf{S}.$$

If $\Phi' \subset \Phi$ has the product property, the *minimum expected average cost* w.r.t. $\Phi'$ is defined as

$$\mathsf{g}_x = \inf_{\phi} \{\mathsf{g}_x(\phi)\}, \quad x \in \mathbf{S}.$$

If for $\phi \in \Phi'$ it holds that $\mathsf{g}(\phi) = \mathsf{g}$, then $\phi$ is called average optimal (in $\Phi'$).

A stronger notion of optimality, called Blackwell optimality, applies more often than is generally noted. We define it next (see also [23]).

Let $\Phi' \subset \Phi$ have the product property. The policy $\phi^* \in \Phi'$ is *Blackwell optimal* w.r.t. $\Phi'$, if for any $x \in \mathbf{S}$, $\phi \in \Phi'$, there exists $\alpha(x,\phi) > 0$, such that $v^\alpha_x(\phi^*) \le v^\alpha_x(\phi)$ for $\alpha < \alpha(x,\phi)$. Additionally, $\phi^*$ is *strongly Blackwell optimal* if $\inf_{x\in\mathbf{S},\phi\in\Phi'} \alpha(x,\phi) > 0$.

## 2.2.1 Discounted cost

To determine the discounted cost $v^\alpha$ an important instrument is the *(discrete time) discount optimality equation* (DDOE)

$$u_x = \inf_{\phi_x\in\Phi_x} \Big\{c_x(\phi_x) + (1-\alpha)\sum_{y\in\mathbf{S}} p_{xy}(\phi_x)u_y\Big\}, \quad x \in \mathbf{S}, \tag{2.2}$$

for $\Phi' = \prod_{x\in\mathbf{S}} \Phi'_x$ having the product property. In this subsection we show that mild conditions guarantee the existence of a unique solution to this equation in a certain space of functions. Moreover, the inf is a min, and a minimising policy in (2.2) is optimal in $\Phi'$ (and even optimal within the larger set of randomised and non-stationary policies generated by $\Phi'$).

The condition used here has been taken from [47, 75].

**Definition 2.2.2.** Let $\gamma \in \mathbb{R}$. The function $V : \mathbf{S} \to (0, \infty)$ is called a $(\gamma, \Phi)$-drift function if $P(\phi)V \leq \gamma V$ for all $\phi \in \Phi$. Note that '$\leq$' stands for component-wise ordering.

**Definition 2.2.3.** The Banach space of $V$-bounded functions on $\mathbf{S}$ is denoted by $\ell^\infty(\mathbf{S}, V)$. This means that $f \in \ell^\infty(\mathbf{S}, V)$ if $f : \mathbf{S} \to \mathbb{R}$ and

$$\|f\|_V = \sup_{x \in \mathbf{S}} \frac{|f_x|}{V_x} < \infty.$$

**Assumption 2.2.2 ($\boldsymbol{\alpha}$).** i) There exist a constant $\gamma < 1/(1-\alpha)$ and a function $V : \mathbf{S} \to (0, \infty)$ such that $V$ is $(\gamma, \Phi)$-drift function and that $\phi \mapsto P(\phi)V$ is component-wise continuous;

ii) $c_V := \sup_\phi \|c(\phi)\|_V < \infty$.

The above assumption allows to rewrite (2.1) as

$$v^\alpha(\phi) = \sum_{n=0}^{\infty} (1-\alpha)^n \, P^{(n)}(\phi) c(\phi).$$

The following lemma is quite straightforward to prove. For completeness we give the details.

**Lemma 2.2.1.** *Suppose that the Assumptions 2.2.1 and 2.2.2 $(\alpha)$ hold, then $\phi \mapsto v^\alpha(\phi)$ is component-wise continuous and $v^\alpha(\phi)$ is the unique solution in $\ell^\infty(\mathbf{S}, V)$ to*

$$u = c(\phi) + (1-\alpha) \, P(\phi) u. \tag{2.3}$$

*Proof.* First notice that $v^\alpha(\phi) \in \ell^\infty(\mathbf{S}, V)$, since

$$\begin{aligned} |v^\alpha(\phi)| = |\sum_{n=0}^{\infty} (1-\alpha)^n \, P^{(n)}(\phi) c(\phi)| &\leq \sum_{n=0}^{\infty} (1-\alpha)^n \, P^{(n)}(\phi) c_V \cdot V \\ &\leq (1-\alpha)^n \gamma^n c_V \cdot V = \frac{c_V}{1-(1-\alpha)\gamma} V. \end{aligned}$$

Next, $v^\alpha(\phi)$ is a solution to Eq. (2.3), since

$$\begin{aligned} (1-\alpha) \, P(\phi) v^\alpha(\phi) &= (1-\alpha) \, P(\phi) \sum_{n=0}^{\infty} (1-\alpha)^n \, P^{(n)}(\phi) c(\phi) \\ &= \sum_{n=1}^{\infty} (1-\alpha)^n \, P^{(n)}(\phi) c(\phi) = v^\alpha(\phi) - c(\phi). \end{aligned}$$

Let $f = (f_x)_x \in \ell^\infty(\mathbf{S}, V)$ be any solution to Eq. (2.3), then

$$\begin{aligned} v_x^\alpha(\phi) - f_x &= (1-\alpha)\sum_y p_{xy}(\phi)(v_y^\alpha(\phi) - f_y) \\ &= (1-\alpha)^n \sum_y p_{xy}^{(n)}(\phi)(v_y^\alpha(\phi) - f_y). \end{aligned}$$

Hence,

$$\begin{aligned} |v_x^\alpha(\phi) - f_x| &\leq (1-\alpha)^n \sum_y p_{xy}^{(n)}(\phi)|v_x^\alpha(\phi) - f_x| \\ &\leq (1-\alpha)^n\, P^{(n)}(\phi)V_x \cdot (c_V + \|f\|_V) \\ &\leq (1-\alpha)^n \gamma^n V_x \cdot (c_V + \|f\|_V) \to 0, \quad n \to \infty. \end{aligned}$$

This implies $f = v^\alpha$, hence $v^\alpha$ is the unique solution to Eq. (2.3) in $\ell^\infty(\mathbf{S}, V)$.

Finally, to show $\phi \mapsto v_x^\alpha(\phi)$, $x \in \mathbf{S}$, is continuous, notice that by assumption $\phi \mapsto P(\phi)V$ is component-wise continuous. It follows that $\phi \mapsto P^{(n)}(\phi)V$ component-wise continuous. Since $P^{(n)}(\phi)V \leq \gamma^n V$, the dominated convergence theorem yields that $\phi \mapsto \sum_{n=0}^\infty (1-\alpha)^n\, P^{(n)}(\phi)V < \infty$ component-wise continuous. Further, since $\phi \mapsto c(\phi)$ is component-wise continuous and $|c(\phi)| \leq c_V \cdot V$, an application of the generalised dominated convergence theorem ([58, Proposition 11.18]) implies component-wise continuity of $\phi \mapsto \sum_{n=0}^\infty (1-\alpha)^n\, P^{(n)}(\phi)c(\phi) = v^\alpha(\phi)$. □

The following theorem is a well-known result by Wessels [75].

**Theorem 2.2.2** (cf. Wessels [75]). *Suppose that $\Phi' = \prod_x \Phi'_x$ has the product property and that Assumptions 2.2.1 and 2.2.2($\alpha$) hold. Then $v^\alpha$ is the unique solution in $\ell^\infty(\mathbf{S}, V)$ to the discounted optimality equation (DDOE) (2.2).*

*Moreover, the infimum is attained as a minimum. For any $\phi^* = (\phi_x^*)_x \in \Phi'$, for which $\phi_x^*$ achieves the minimum in Eq. (2.2) for all $x \in \mathbf{S}$, it holds that $v^\alpha(\phi^*) = v^\alpha$ and $\phi^*$ is ($\alpha$-discounted) optimal in $\Phi'$.*

The versatile applicability of $(\gamma, \Phi)$-drift functions is illustrated in Chapters 4 and 5, and the example below.

**Example 2.2.1.** First note for the bounded cost case, that the function $V_x \equiv 1$ is an appropriate function satisfying Assumption 2.2.2($\alpha$).

As a simple example, consider a discrete time single server queue, where the probability of an arrival in the next time slot is $p \in (0, 1)$. The system state represents the number of customers in the system, hence, the state space is

$\mathbf{S} = \{0, 1, \ldots\}$. The probability of a service completion in the next time slot depends on a parameter $\phi = \{\phi_x\}_{x=1}^{\infty}$, where $\phi_x$ stands for the probability of a service completion in the next time-slot, independent of arrivals, when the system state is $x$. The parameter $\phi$ stands for an a priori determined service control. The parameter space $\Phi$ maybe any compact subset of $\{0\} \times [0,1]^{\infty}$.

To any fixed parameter $\phi$, one may associate a Markov process, representing the system state at any time $t$, with transition probabilities given by

$$p_{xy}(\phi) = \begin{cases} p(1-\phi_x), & y = x+1 \\ (1-p)\phi_x, & y = (x-1)\mathbb{1}_{\{x \neq 0\}} \\ 1 - p - \phi_x + 2p\phi_x, & y = x. \end{cases}$$

As an appropriate $(\gamma, \Phi)$-drift function, we may choose $V_x = e^{\epsilon x}$, with $\epsilon > 0$ to be determined below:

$$\begin{aligned} \sum_y p_{xy}(\phi) e^{\epsilon y} &= (1-p)\phi_x e^{\epsilon(x-1)} + (1 - (1-p)\phi_x - p(1-\phi_x))e^{\epsilon x} \\ &\quad + p(1-\phi_x)e^{\epsilon(x-1)} \\ &= e^{\epsilon x}\Big(1 + p(1-\phi_x)(e^{\epsilon} - 1) + (1-p)\phi_x(1 - e^{-\epsilon})\Big) \\ &\leq e^{\epsilon x}\big(1 + p(e^{\epsilon} - 1)\big). \end{aligned}$$

For $\epsilon = 0$, the coefficient of $e^{\epsilon x}$ in the above equals 1. Since $1/(1-\alpha) > 1$, one can always choose $\epsilon$ small enough so that

$$\gamma := e^{\epsilon x}\big(1 + p(e^{\epsilon} - 1)\big) < \frac{1}{1-\alpha}.$$

As a consequence Assumption 2.2.2($\alpha$),($i$) is satisfied. The example shows, the existence of a $(\gamma, \Phi)$-drift function does not impose any restrictions on the class structure of the associated Markov processes and transience is allowed as well. Moreover, it is often a good and simply checked choice to take $V$ exponential. Since generally cost structures are linear or quadratic as a function of state, they are dominated by exponential functions. Thus, they fit in the framework discussed here.

**Value Iteration** A very important algorithm to calculate $v^\alpha$ is the value iteration algorithm (VI), originally due to Bellman [9].

**Algorithm 1** VI for an $\alpha$-discounted cost $\epsilon$-optimal policy

1. Select $v^{\alpha,0} \in \ell^\infty(\mathbf{S}, V)$, specify $\epsilon > 0$, set $n = 0$.
2. For each $x \in \mathbf{S}$, compute $v_x^{\alpha,n+1}$ by

$$v_x^{\alpha,n+1} = \min_{\phi_x \in \Phi'_x} \Big\{ c_x(\phi_x) + (1-\alpha) \sum_{y \in \mathbf{S}} p_{xy}(\phi_x) v_x^{\alpha,n} \Big\}, \tag{2.4}$$

   and let

$$\phi^{n+1} \in \arg\min_{\phi \in \Phi'} \{ c(\phi) + (1-\alpha)\, P(\phi) v^{\alpha,n} \}.$$

3. If

$$\|v^{\alpha,n+1} - v^{\alpha,n}\|_V \le \frac{1-(1-\alpha)\gamma}{2(1-\alpha)\gamma}\epsilon,$$

   then put $v^\epsilon := v^{\alpha,n+1}$, $\phi^\epsilon := \phi^{n+1}$, stop. Otherwise increment $n$ by 1 and return to step 2.

---

**Theorem 2.2.3** (cf. [75], [54, Theorem 6.3.1])**.** *Suppose that $\Phi' = \prod_x \Phi'_x \subset \Phi$ has the product property and that Assumptions 2.2.1 and 2.2.2($\alpha$) hold. Let let $v^{\alpha,0} \in \ell^\infty(\mathbf{S}, V)$ and $\epsilon > 0$. Let $\{v^{\alpha,n}\}_{n\in\mathbb{N}}$ satisfy Eq. (2.4) for $n \ge 1$. Then the following hold.*

**i)** $\lim_{n\to\infty} \|v^\alpha - v^{\alpha,n}\|_V = 0$*, in particular,*

$$\begin{aligned} \|v^\alpha - v^{\alpha,n}\|_V &\le \frac{1}{1-(1-\alpha)\gamma} \|v^{\alpha,n+1} - v^{\alpha,n}\|_V \\ &\le \frac{((1-\alpha)\gamma)^n}{1-(1-\alpha)\gamma} \|v^{\alpha,1} - v^{\alpha,0}\|_V. \end{aligned}$$

*Any limit point of the sequence $\{\phi^n\}_n$ is an $\alpha$-discount optimal policy.*

**ii)** *$v^\epsilon$ is an $\epsilon/2$-approximation of $v^\alpha$, in other words, $\|v^\alpha - v^{\alpha,n+1}\|_V \le \frac{\epsilon}{2}$.*

**iii)** *$\phi^\epsilon$ is an $\epsilon$-optimal policy, in other words, $\|v^\alpha - v^\alpha(\phi^\epsilon)\|_V \le \epsilon$.*

*Proof.* The proof of Theorem 2.2.3 (i) is straightforward using that

$$v^{\alpha} - v^{\alpha,n} = \lim_{N\to\infty} \sum_{k=n}^{N} (v^{\alpha,k+1} - v^{\alpha,k}).$$

The bounds are somewhat implicit in [75]. They are completely analogously to the bounds of e.g. [54, Theorem 6.3.1] for the bounded reward case, with $\lambda$ replaced by $(1-\alpha)\gamma$. The derivation is similar. □

*Remark* 2.2.2. The reader may wish to point out that a solution to the $\alpha$-DDOE yielding an $\alpha$-discount deterministic policy exists without any further conditions in the case of non-negative cost (negative dynamic programming, cf. [70]) and a finite action space per state. Also VI converges provided $v_0 \equiv 0$, although no convergence bounds can be provided. If the action space is compact, additional continuity and inf-compactness (cf. [30, Corollary 5.7]) properties are necessary for the existence of a stationary deterministic policy attaining the minimum in the $\alpha$-DDOE. It is not clear to us how these conditions could be extended in order to include parametrised Markov processes.

Notice further, that unfortunately in general there is no unique solution to the $\alpha$-DDOE (cf. [30], [63, Section 4.2]). Using norm conditions as in this chapter, allows to identify the value function as the unique one in the Banach space of functions bounded by $V$ (cf. Theorem 2.2.2). In the non-negative cost case, the value function is the minimum solution to the $\alpha$-DDOE (see [63, Theorem 4.1.4]).

In case of a finite state space, VI can be numerically implemented. In the case of a countable space, its use is restricted to the derivation of structural properties of the value function and $\alpha$-discount optimal policy. Structural properties such as non-decreasingness, convexity, etc. can be used to show for instance that a threshold policy or an index policy is optimal.

To prove properties via VI, first select a function $v_0$ possessing the properties of interest. Then show by induction that $v^{\alpha,n}$ has this property for all $n$. Under the assumptions of Theorem 2.2.3 one has $v^{\alpha,n} \to v^{\alpha}$, for $n \to \infty$, and so we may conclude $v^{\alpha}$ has this property as well. The existence of an optimal policy with desired properties can be directly derived from the structure of the value function $v^{\alpha}$ in combination with the $\alpha$-DDOE. Alternatively, this can be deduced from the fact that since each $\phi^n$ has these properties, any limit point has.

The main reference on the propagation of structural properties through the VI induction step Eq. (2.4) is [45]. The technique discussed in this monograph is called *event based dynamic programming*, and it presents a systematic framework of propagations of the desired structural properties for operators that represent events. We have developed new operators in [11, 16], as well as in Chapters 4, 5 and 7, for special perturbations or truncations of non-uniformisable MDPs, as described below. In Section 2.3.4 we present an example.

## 2.2.2 Approximations/Perturbations

Next we focus our attention to parameters capturing a perturbation of the MDP. This parameter set should capture the collection of deterministic policies $\mathcal{D}$, as well as a perturbation set $\mathcal{N}$. This perturbation can have multiple interpretations, depending on the context. It can be a finite state approximation, or it can represent some uncertainty in the input parameters. Put $\Phi = \mathcal{N} \times \mathcal{D}$. Notice, that the set $\{N\} \times \mathcal{D} \subset \Phi$ need not automatically have the product property, $N \in \mathcal{N}$.

The following continuity result follows directly from Lemma 2.2.1.

**Corollary 2.2.4** (to Lemma 2.2.1 and Theorem 2.2.2)**.** *Suppose that Assumptions 2.2.1, and 2.2.2($\alpha$) hold. Further assume that $\{N\} \times \mathcal{D}$ has the product property, for $N \in \mathcal{N}$. Then,*

*i)* $\lim_{N \to N_0} v^\alpha(N) = v^\alpha(N_0)$*;*

*ii) any limit point of $\{\delta^*_N\}_{N \to N_0}$ is optimal in $\{N_0\} \times \mathcal{D}$.*

Without the existence of a $(\gamma, \Phi)$-drift function bounding the one-step cost uniformly in the parameter, the above convergence result may fail to hold.

**Example 2.2.2.** (cf. [63, Example 4.6.1]) Let the parameter set be given by $\mathcal{N} = \{3, 4, \ldots, \infty\}$, and state space $\mathbf{S} = \{0, 1, \ldots\}$ for $N \in \mathcal{N}$. The transition probabilities are as follows.

$$p_{xy}(\infty) = \frac{1}{2} \quad y \in \{0, x+1\},$$

and for $N < \infty$

$$p_{xy}(N) = \begin{cases} \frac{1}{2}, & x \neq N-1, N, y = 0 \\ \frac{1}{2} - \frac{1}{N}, & x \neq N-1, N, y = x+1 \\ \frac{1}{N}, & x \neq N, y = N \\ 1 - \frac{1}{N}, & x = N-1, y = 0 \\ 1, & x = y = N. \end{cases}$$

Further let $\alpha < \frac{1}{2}$ and define $c_x(N) = x^2$ for $N \leq \infty$. The calculations in [63, Example 4.6.1] show that $v_0^\alpha(\infty) < \lim_{N\to\infty} v_0^\alpha(N) = \infty$. It is simply checked that any $(\gamma, \Phi)$-drift function can at most be linear. Indeed for $1 < N < \infty$, it must hold that

$$\frac{1}{2}V_0 + (\frac{1}{2} - \frac{1}{N})V_2 + \frac{1}{N}V_N \leq \gamma V_1,$$

leading to the requirement that $V_N \leq N\gamma V_1$, hence $\sup_{N,x} \frac{c_x(N)}{V_x} = \infty$.

Literature related to Corollary 2.2.4 can be found in [54, Section 6.10.2] and [63, Section 4.6]. Both consider finite space truncations. The first reference further assumes the existence of a $(\gamma, \Phi)$-drift function without continuity, but with an additional tail condition and a prescribed truncation method. These conditions are implied by ours.

The second reference considers minimisation of non-negative costs, as well as a finite action space per state. No assumption on the truncation method needs to be made if there is a uniform bound on the costs. If the costs are allowed to be unbounded as a function of state, then conditions on the truncation method have to be made. A related setup to the one presented in the present paper is discussed in [52, Theorem 3.1], but the conditions imposed are (slightly) more restrictive (cf. Remark 3.5.2).

**Example 2.2.3.** The above example is a case where neither the conditions from [63] are met, nor the ones presented in this paper.

However, the conditions in the approximating sequence method of [63] are not even fulfilled, if we change the cost function to $c_x(N) = x$ for all $N \leq \infty$, $x \in \mathbf{S}$. On the other hand, the function $V$ defined by $V_x = x$, $x \in \mathbf{S}$, is a $(\gamma, \mathcal{N})$-drift function, for which the conditions of Corollary 2.2.4 are trivially met. Hence, the set-up in this paper can be applied and we find that $\lim_{N\to\infty} v_0^\alpha(N) = v_0^\alpha(\infty) < \infty$.

**Type of perturbations** Although any perturbation satisfying the conditions of Corollary 2.2.4 yields (component-wise) continuity of the value function as a function the perturbation parameter, not any perturbation is desirable in terms of structural properties.

To explain this, consider the following server allocation problem, see Figure 2.1. Customers arrive at a service unit according to a Poisson($\lambda$) process. Their service time at unit 1 takes an exponentially distributed amount of time with parameter $\mu_1$. After finishing service in unit 1, with probability $p_1$ an

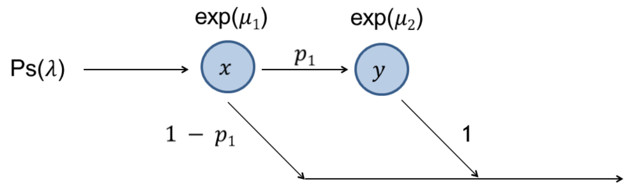



Figure 2.1: Tandem queue

additional $\exp(\mu_2)$ amount of service is requested in unit 2, and with probability $1-p_1$ the customer leaves the network, $p_1 \in (0,1)$. There is only one server who has to be allocated to one of the units. We assume that idling is not allowed. The goal is to determine an allocation policy that minimises the $\alpha$-discounted holding cost. The holding cost per unit time is given by the number of customers in the system.

Clearly this problem can be modelled as a continuous time MDP. However, it is equivalent to study the associated uniformised discrete time system (cf. Section 2.3.1). The data of this discrete time MDP are as follows. Denote by $X_n$ the number of customers in unit 1 and unit 2 respectively, at time $n$, $n = 0, 1, \ldots$. Then $\mathbf{S} = \mathbf{Z}_+^2$, where state $(x_1, x_2)$ represents that $x_1$ customers are present in unit 1 and $x_2$ in unit 2. By uniformisation we may assume that $\lambda + \mu_1 + \mu_2 = 1$, and so the rates represent probabilities. Independently of the allocation decision, a cost $c_x(\phi) = x_1 + x_2$ is incurred.

Suppose that the state equals $x$. If both units are non-empty, then either unit 1 is served (decision $\delta_x = 1$) or unit 2 (decision $\delta_x = 2$). If one of the units is empty, but not both, the server will be allocated to the non-empty unit during the next time-slot. This leads to the following transition probabilities:

$$p_{xy}(\delta_x) = \begin{cases} \lambda & y = (x_1+1, x_2) \\ p_1\mu_1 & x_1 > 0, y = (x_1-1, x_2+1), \delta_x = 1 \\ (1-p_1)\mu_1 & x_1 > 0, y = (x_1-1, x_2), \delta_x = 1 \\ \mu_2 & x_2 > 0, y = (x_1, x_2-1), \delta_x = 2 \\ 1 - \sum\limits_{w \neq x} p_{xw}(\delta_x) & y = x. \end{cases}$$

Let the discount factor $\alpha$ be given. It is easily verified that there exists a

$(\gamma, \mathcal{D})$-drift function $V : \mathbf{S} \to \mathbb{R}_+$ of the form

$$V(x, y) = e^{\epsilon_1 x_1 + \epsilon_2 x_2}, \tag{2.5}$$

with $\gamma < 1/(1-\alpha)$ and $\epsilon_1, \epsilon_2 > 0$. Assumptions 2.2.1 and 2.2.2 ($\alpha$) are satisfied for $V$ and $\Phi = \mathcal{D}$ and so the results of Theorem 2.2.2 apply.

Assume that $(1 - p_1)\mu_1 > \mu_2$. By using VI, event based dynamic programming yields that allocating the server to unit 1, when non-empty, is $\alpha$-discount optimal. Indeed, this gives a larger cost reduction per unit time due to customer service completions than allocating to unit 2. Thus, noticing that the cost rate per unit time and per server unit are equal to 1, this allocation policy is a generalised $c\mu$-rule. Let us refer to this policy as AP1 (allocation to unit 1 policy).

Since this is true for any $0 < \alpha < 1$, AP1 is *strongly Blackwell optimal.* We therefore expect to see this structure in numerical experiments. To perform such an experiment, one needs truncate the state space.

**Straightforward perturbation** A straightforward truncation can be as follows. Choose $M, N \in \mathbf{Z}_+$, $N, M > 0$. At the truncation boundary $\{x \mid x_1 = N, \text{ and/or } x_2 = M\}$, transitions leading out of the rectangle $\{x \mid x_1 \leq N, x_2 \leq M\}$ are redirected back as follows. New arrivals in states $\{x | x_1 = N\}$ are redirected to the same state. A service completion of a type 1 customer in states $\{x | x_2 = M\}$ has a success probability of 1, that is, it remains in states with $x_2 \leq M$. The perturbation set is $\mathcal{N} = \big\{(N, M) \mid N, M \in \{1, 2, \ldots\}\big\}$. Also in this case, one can easily check that there exist $\epsilon_1$, $\epsilon_2 > 0$ and $\gamma \in \mathbb{R}$, such that $V$ from (2.5) is a $(\gamma, \mathcal{N} \times \mathcal{D})$-drift function for $\epsilon_1, \epsilon_2$ small enough. Moreover, Assumptions 2.2.1 and 2.2.2($\alpha$) are satisfied for this $V$ and for $\Phi = \mathcal{N} \times \mathcal{D}$. Note that the states outside the rectangle $\{x \mid x_1 \leq N, x_2 \leq M\}$ have become transient, and so the choice of actions in those states does not affect the optimal actions within the rectangle.

Let $\delta^*$ be $\alpha$-discount optimal for $\mathcal{D}$, and $\delta^{N,M}$ for $\{(N, M)\} \times \mathcal{D}$. Then by virtue of Lemma 2.2.1

$$v^\alpha(\delta^{N,M}) \to v^\alpha(\delta^*), \quad (N, M) \to \infty,$$

and any limit policy is optimal. However, choosing the following parameters: $\lambda = 0.102$, $\mu_1 = 0.870$, $\mu_2 = 0.028$, $p_1 = 0.22$ and $N = M = 300$ leads to the optimal policy in the rectangle shown in the picture below. The grey color stands for server allocation to unit 2.

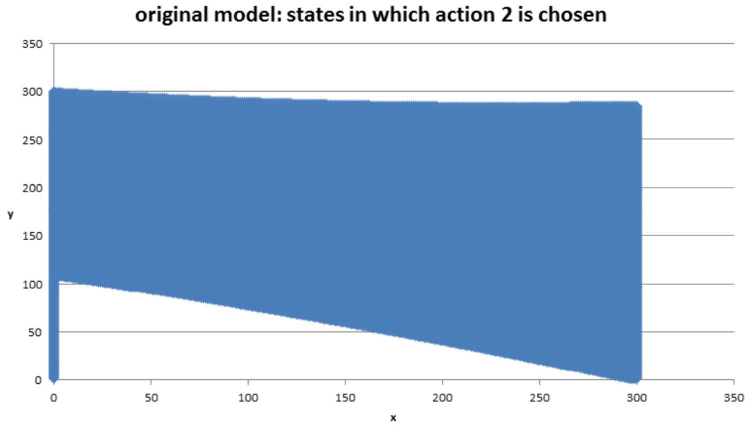



Figure 2.2: Standard Truncation

This optimal policy is very far from being the index policy AP1, although the truncation size seems large enough to exhibit an optimal policy that is 'closer' to AP1. One starts to wonder what the effect of such a straightforward truncation has been on numerical approximations of other models studied in the literature.

**Smoothed rate truncation (SRT)** SRT is a perturbation introduced in [11], where 'outward bound' probabilities are linearly decreased as a function of state. This creates a perturbed MDP with a finite closed class under any policy. It is not meaningful to try and give a complete definition, but the idea is best illustrated by specifying possible SRT's for the above example. The experience with SRT so far is that it leaves the structure of an optimal policy intact (cf. [11] and 5). On the other hand, since it perturbs transition probabilities from *all* states in the finite closed class, the value function itself seems better approximated by a straightforward cut-off, such as the one as described above.

One can apply SRT as follows. Fix $N, M$. Then we put

$$p_{xy}(N,M,\delta_x) = \begin{cases} \lambda(1-\frac{x_1}{N})^+ & y=(x_1+1,x_2) \\ p_1\mu_1\big(1-\frac{x_2}{M}\big)^+ & x_1>0, y=(x_1-1,x_2+1), \delta_x=1 \\ (1-p_1)\mu_1 + & \\ \quad p_1\mu\frac{x_2}{M}\mathbb{1}_{\{x_2\le M\}} & x_1>0, y=(x_1-1,x_2), \delta_x=1 \\ \mu_2 & x_2>0, y=(x_1,x_2-1), \delta_x=2 \\ 1-\sum_{w\neq x} p_{xw}(N,M,\delta_x) & y=x. \end{cases}$$

Again, the function $V$ from (2.5) is a $(\gamma, \mathcal{N} \times \mathcal{D})$-drift function satisfying Assumptions 2.2.1 and 2.2.2($\alpha$).

The following picture illustrates the numerical results with $N = M = 35$. This confirms the results in [11] and Chapter 5, suggesting that SRT allows to obtain information on the structure of an optimal policy for an infinite state MDP.

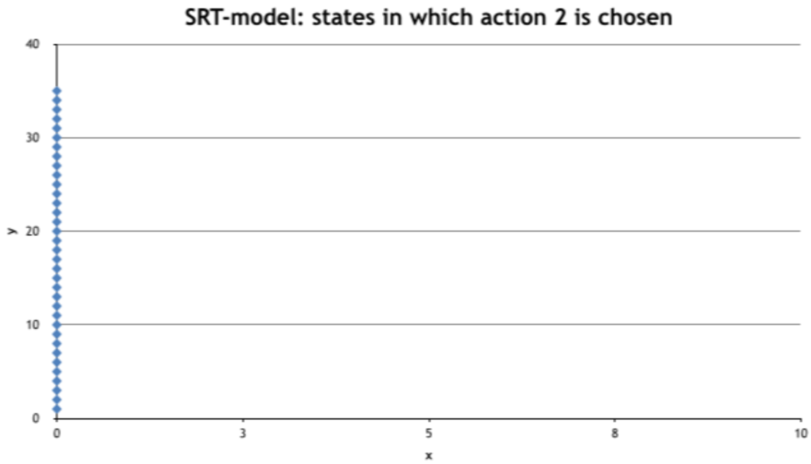



Figure 2.3: Smoothed Rate Trunction

Now in both the truncated and the non-truncated MDP AP1 is optimal. We have not proven this result, but we did so for another SRT (cf. Chapter 6) that is not based on a rectangle but on a triangle. Using a triangular truncation requires less events to be truncated and so the proofs are simpler. Notice, that the above also illustrates that smooth truncations are not uniquely defined, but different choices may be possible.

Fix $N$. Then we put

$$p_{xy}(N, \delta_x) = \begin{cases} \lambda(1 - \frac{x_1+x_2}{N})^+, & y = (x_1+1, x_2) \\ p_1\mu_1 & x_1 > 0, y = (x_1 - 1, x_2 + 1), \delta_x = 1 \\ (1-p_1)\mu_1 & x_1 > 0, y = (x_1 - 1, x_2), \delta_x = 1 \\ \mu_2 & x_2 > 0, y = (x_1, x_2 - 1), \delta_x = 2 \\ 1 - \sum_{w \neq x} p_{xw}(N, \delta_x) & y = x. \end{cases}$$

Using event based dynamic programming (with special SRT operators), Corollary 6.3.5 shows that AP1 is optimal, for each $N \in \mathbb{N}$.

### 2.2.3 Average cost

Establishing a framework for the average cost optimality criterion is more difficult than for the discounted cost case. There are several cautionary examples in the literature highlighting the complications. In our opinion, the most intricate one is the Fisher-Ross example [31]. In this example, all action spaces are finite, the Markov process associated with any stationary deterministic policy is irreducible and has a stationary distribution. However, there is an optimal non-stationary policy, but no stationary deterministic one is optimal.

In this chapter we provide conditions under which there exists a stationary average optimal policy satisfying the average cost optimality equation. The proof requires the vanishing discount approach. Our focus in this chapter are non-negative cost functions, the analysis of which does not require a heavy drift condition that imposes positive recurrence of the Markov process associated with any parameter. However, below, we do touch upon benefits of using the heavier conditions.

The conditions that we will focus on, are based on the work of Borkar [21], in the form discussed in [63]. They resemble the conditions from the earlier work in [74]. They imply conditions developed by Sennott in [61], requiring (i) lower-bounded direct cost; (ii) the existence of a finite $\alpha$-discounted value function, (iii) the existence of a constant $L$ and a function $M : \mathbf{S} \to \mathbb{R}$, such that $-L \leq v^\alpha_x - v^\alpha_z \leq M_x$ for some state $z$, and all $\alpha$ sufficiently small, and (iv) for each $x \in \mathbf{S}$, there exists $\phi_x$, such that $\sum_y p_{xy}(\phi_x) M_y < \infty$. Under these conditions Sennott [61] proves the statement in Theorem 2.2.5 below, with equality in (2.6) replaced by inequality, and without property (3) from Theorem 2.2.5. It is appropriate to point out that the approach initiated in [61] was based on a bounded cost average optimality result in [56].

The following definitions are useful. Define $\tau_z := \min_{n\geq 1} \mathbb{1}_{\{z\}}(X_n)$ to denote the hitting time of $z \in \mathbf{S}$. Let $m_{xz}(\phi) = \mathsf{E}^\phi_x\big[\tau_z\big]$ and $c_{xz}(\phi) = \mathsf{E}^\phi_x\big[\sum_{n=0}^{\tau_z} c_{X_n}(\phi)\big]$.

**Assumption 2.2.3.** Let $\Phi' = \prod_x \Phi'_x \subset \Phi$ have the product property. The following holds.

i) Non-negative cost rates: $c_x(\phi) \geq 0$ for all $x \in \mathbf{S}, \phi \in \Phi'$.

ii) There exist $z \in \mathbf{S}$ and $\phi_0 \in \Phi'$ such that $m_{xz}(\phi_0), c_{xz}(\phi_0) < \infty$ for all $x \in \mathbf{S}$, with the potential exception of $x = z$, that is allowed to be absorbing. Note that this implies that $\mathsf{g}_x(\phi_0)$ is independent of $x \in \mathbf{S}$ and hence we write $\mathsf{g}(\phi_0) = \mathsf{g}_x(\phi_0)$, for all $x \in \mathbf{S}$.

iii) There exists $\epsilon > 0$ such that $D = \{x \in \mathbf{S} \mid c_x(\phi) \leq \mathsf{g}(\phi_0) + \epsilon$ for some $\phi \in \Phi'\}$ is a finite set.

iv) For all $x \in D$ there exists $\phi^x \in \Phi'$ such that $m_{zx}(\phi^x), c_{zx}(\phi^x) < \infty$.

**Theorem 2.2.5.** *Suppose that Assumptions 2.2.1, 2.2.2($\alpha$), for $\alpha \in (0,1)$, and 2.2.3 hold. Then the following holds.*

**i)** *There exists a solution tuple $(g^*, v^*)$, $g^* \in \mathbb{R}_+$, $v^* : \mathbf{S} \mapsto \mathbb{R}$, to the average cost optimality equation (DAOE)*

$$g + u_x = \min_{\phi_x \in \Phi'_x} \Big\{ c_x(\phi_x) + \sum_{y \in \mathbf{S}} p_{xy}(\phi_x) u_y \Big\}, \tag{2.6}$$

*with the property that (1) $g^* = \mathsf{g}$ is the minimum expected average cost (in $\Phi'$), (2) any $\phi^* \in \Phi'$ with*

$$\phi^*_x \in \arg\min_{\phi_x \in \Phi'_x} \Big\{ c_x(\phi_x) + \sum_{y \in \mathbf{S}} p_{xy}(\phi_x) v_y \Big\}$$

*is (average cost) optimal in $\Phi'$ and (3) there exists $x^* \in D$ with $v^*_{x^*} = \inf_x v^*_x$.*

**ii)** *Let $x_0 \in \mathbf{S}$. Any sequence $\{\alpha_n\}_n$ with $\lim_n \alpha_n = 0$, has a subsequence, again denoted $\{\alpha_n\}_n$, along which the following limits exist:*

$$\begin{aligned}
v'_x &= \lim_{n\to\infty} (v^{\alpha_n}_x - v^{\alpha_n}_{x_0}), \quad x \in \mathbf{S}, \\
g' &= \lim_{n\to\infty} \alpha_n v^{\alpha_n}_x, \quad x \in \mathbf{S} \\
\phi' &= \lim_{n\to\infty} \phi^{\alpha_n}.
\end{aligned}$$

*Further, the tuple $(g', v')$ is a solution to Eq. (2.6) with the properties (1), (2) and (3), so that $g' = \mathsf{g}$. Moreover, $\phi'$ takes minimising actions in Eq. (2.6) for $g = g'$ and $u = v'$.*

Theorem 2.2.5 is a slight extension from [63, Theorem 7.5.6], where the action space is assumed to be finite. Although the various necessary proof parts are scattered over [63, Chapter 7], we will merely indicate the necessary adjustments to allow for the compact parameter case. We would like to note that we use a completely analogous reasoning in the proof of Theorem 2.3.4, which contains all further details.

*Proof.* A close examination of the proof of [63, Theorem 7.5.6] shows that the assumption of a finite action space is not necessary. The proof can be adjusted in such a way that the statement holds for a compact action space as well. We briefly discuss the adjustments below. The existence of the limits along a sequence $\{\alpha_n\}_n$, $\alpha_n \downarrow 0$, $n \to \infty$, in assertion *ii)* is a direct result of Sennott [63].

Obtaining the average cost optimality inequality (DAOI) for a limit point of $\alpha$-discount optimal policies, as $\alpha \to 0$, can be achieved by virtue of Fatou's lemma. This policy is shown to be optimal.

Further, one needs to show explicitly that there exists a policy realising the infimum of Eq. (2.6). Since the limit policy satisfies the DAOI, a similar (very ingenious) reasoning as in the proof of Sennott [63, Theorem 7.4.3] yields that this policy satisfies the DAOE as well. In fact, any policy satisfying the DAOI also satisfies the DAOE. It can then be shown by contradiction that this limit policy must attain the infimum. As a consequence, the limit tuple $(g', v')$ from (ii) is a solution to Eq. (2.6). The rest directly follows from the proof of the afore mentioned theorem in [63]. $\square$

*Remark* 2.2.3. In the literature the formulation of statements similar to Theorem 2.2.5 on the DAOE may sometimes have a misleading character. This may occur when the existence of a solution to Eq. (2.6) is stated first, and a subsequent claim is made that any minimising policy in Eq. (2.6) is average cost optimal. Strictly speaking, this may not be true. Examples 2.2.4 and 2.2.5 below, at the end of this section, illustrate that other 'wrong' solutions may exist. Unfortunately, Assumption 2.2.3 does not admit tools to select the 'right' solution among the set of all solutions. Thus, under Assumption 2.2.3 a solution to the DAOE should always be obtained via the vanishing discount approach, as in Theorem 2.2.5 (ii).

The next issue to be discussed is how to verify Assumption 2.2.3 (ii) and Assumption 2.2.3 (iv). This can be inferred from the following Lyapunov function criterion, which is a direct application of [40, Lemma 3.1]. The proof is a simple iteration argument.

**Lemma 2.2.6.** *Let $x_0 \in \mathbf{S}$ be given. Let $\phi \in \Phi$. Suppose that there exist functions $f, h : \mathbf{S} \to [0, \infty)$ with*

**i)** $f_x \geq \max\{1, c_x(\phi)\}$, $x \in \mathbf{S} \setminus \{x_0\}$;

**ii)** $f_x + \sum_{y \neq x_0} p_{xy}(\phi) h_y \leq h_x$, $x \in \mathbf{S}$.

*Then* $m_{xx_0}(\phi), c_{xx_0}(\phi) \leq h_x$, $x \in \mathbf{S}$.

To pave the way for developing a roadmap for obtaining structures of average cost optimal policies, we will shortly discuss the applicability of VI. Let us first state the algorithm. Again assume that $\Phi' \subset \Phi$ has the product property.

**Algorithm 2** VI for an expected average cost optimal policy

1. Select $v^0$, set $n = 0$.
2. For each $x \in \mathbf{S}$, compute $v_x^{n+1}$ by
$$v_x^{n+1} = \min_{\phi_x \in \Phi'_x} \Big\{ c_x(\phi_x) + \sum_{y \in \mathbf{S}} p_{xy}(\phi_x) v_y^n \Big\},$$
and let
$$\phi^{n+1} \in \arg\min_{\phi \in \Phi'} \{ c(\phi) + P(\phi) v^n \}.$$
3. Increment $n$ by 1 and return to step 2.

To our knowledge there are relatively few non-problem specific papers on the convergence of average cost VI for countable state space MDPs, cf. [41], [62], [6], and [2], the latter of which is based on the thesis [65]. The conditions in the first three papers are not restricted to conditions on the input parameters. In our opinion, the easiest verifiable ones are contained in the paper [6], involving properties of the set of policies $\{\phi_n\}_n$. In case of well-structured problems, say $\phi_n$ are all equal, or have very specific structures, these conditions are easy to verify.[1] Here, we will restrict to the conditions from [65] and [2] that are, as far as we know, the only ones formulated directly in terms of the input parameters of the process. The notation $\mathbf{e}$ stands for the function on $\mathbf{S}$ identically equal to 1.

**Theorem 2.2.7.** *Let $\Phi'$ have the product property. Suppose that the following drift condition, called $V$-geometric recurrence, holds: there exist a function $V : \mathbf{S} \to [1, \infty)$, a finite set $M \subset \mathbf{S}$ and a constant $\beta < 1$, such that*
$$\sum_{y \notin M} p_{xy}(\phi) V_y \le \beta V_x, \quad x \in \mathbf{S}, \phi \in \Phi'.$$
*Suppose further that the following holds as well:*

[1] we still have a suspicion that there is a gap in the proofs in [6]

- *Assumption 2.2.1;*
- $\sup_{\phi\in\Phi'} \|c(\phi)\|_V < \infty$*;*
- $\phi \mapsto P(\phi)V$ *is component-wise continuous on* $\Phi'$*;*
- *the Markov process with transition matrix* $P(\phi)$ *is aperiodic and has one closed class,* $\phi \in \Phi'$*.*

*Let* $0 \in \mathbf{S}$*. There is a unique solution pair* $(g^*, v^*)$ *with* $v^* \in \ell^\infty(\mathbf{S}, V)$*, and* $v^*_0 = 0$*, to Eq. (2.6) with the properties in Theorem 2.2.5.*

*Furthermore, average cost VI converges, that is,* $\lim_{n\to\infty}(v^n - v^n_0\mathbf{e}) \to v^*$*, and any limit point of the sequence* $\{\phi^n\}_n$ *is average cost optimal and a minimising policy in the DAOE (2.6) with solution tuple* $(g^*, v^*)$*.*

The $V$-uniform geometric recurrence condition in Theorem 2.2.7 has been introduced in [24], and shown in [25] to imply the assertion in Theorem 2.2.7. The paper [25], see also [65], has derived an equivalence of this condition (under extra continuity conditions) with $V$-uniform geometric ergodicity. The thesis [65] additionally shows a similar implication for bounded jump Markov decision processes in continuous time, by uniformisation. Both properties have been extensively used both in the case of a parameter space consisting of one element only (cf. [49] and later works), and in the case of product parameter spaces in the context of optimal control. Together with the negative dynamic programming conditions developed by [61], the $V$-uniform geometric recurrence and ergodicity, developed in [23] and [24], have become 'standard' conditions in many papers and books. See for instance [36], [35], and [53], as well as references therein, for a survey and two books using both types of conditions. A parametrised version of [25] in both discrete and continuous time is currently in preparation. The drawback of using $V$-geometric recurrence is that it implies that each associated Markov process is positive recurrent. This is a major disadvantage for many models and therefore our motivation for using Assumption 2.2.3. Note that customer abandonment has a strong stabilising effect on the associated Markov processes, and then $V$-geometric recurrence typically may apply.

**Roadmap to structural properties** Below we formulate a scheme for deriving the structure of an optimal policy and value function, if the optimisation criterion is to minimise the expected average cost. Let $\Phi' = \Phi = \mathcal{D}$ be the set of all stationary, deterministic policies.

**Roadmap for average cost MDPs in discrete time**

1. Check the conditions of Theorem 2.2.7.
   If satisfied then:
   - perform VI Algorithm 2.
2. If not satisfied, then check Assumptions 2.2.1, 2.2.2 ($\alpha$), for all $\alpha \in (0, 1)$, and 2.2.3. If satisfied then:
   **a)** perform VI Algorithm 1 for the $\alpha$-discounted cost criterion. If there exists $\alpha_0 > 0$ such that the desired structural properties hold for all $\alpha \in (0, \alpha_0)$ then
   **b)** apply the vanishing discount approach by taking the limit $\alpha \to 0$. This is justified by Theorem 2.2.5.
3. If not satisfied, or if no structural properties are concluded, then the outcome is inconclusive.

Note that the vanishing discount approach has the advantage of allowing a conclusion on Blackwell optimality of the limiting policy. Next we provide examples showing that the DAOE may have more than one solution.

**Example 2.2.4.** Consider a simple random walk on the state space $\mathbf{S} = \mathbf{Z}$ without any control. Thus $\Phi$ consists of one element $\phi$, say $\phi_x = 1$ for all $x \in \mathbf{S}$. The transition mechanism is given by

$$p_{xy}(1) = \begin{cases} \lambda, & y = x + 1 \\ \mu, & y = x - 1, \end{cases}$$

where $\lambda < \mu$ and $\lambda + \mu = 1$. The cost in state $x \neq 0$ is equal to $c_x = 1$, and $c_0 = 0$. This is a transient Markov process, and hence it does not satisfy Assumption 2.2.3. However, it does satisfy the assumptions of [61], implying the assertion of Theorem 2.2.5 to hold.

This implies that the vanishing discount approach yields solution tuple $g = 1$ and

$$v_x = \begin{cases} \frac{1-(\mu/\lambda)^x}{\mu-\lambda}, & x < 0 \\ 0, & x \geq 0, \end{cases}$$

if $x_0 = 0$ is chosen. This can be deduced from boundedness conditions that will be discussed in a forthcoming paper [69].

However, other solutions $(g = 1, v')$ exist, namely for any $\theta \in \mathbb{R}$

$$v'_x = \begin{cases} (1-\theta)\frac{1-(\mu/\lambda)^x}{\mu-\lambda}, & x < 0 \\ 0, & x = 0 \\ \theta\frac{(\mu/\lambda)^x-1}{\mu-\lambda}, & x > 0. \end{cases}$$

There is no a priori tool to determine which solution is the one obtained from the vanishing discount approach.

**Example 2.2.5.** Next we restrict the simple random walk to $\mathbf{S} = \mathbf{Z}_+$, and associate the corresponding transitions with $\phi^1$. In other words, putting $\phi^1_x = 1$, $x \in \mathbf{S}$, gives

$$p_{xy}(1) = \begin{cases} \lambda, & y = x+1 \\ \mu, & y = (x-1)^+, \end{cases}$$

where $\lambda < \mu$ and $\lambda + \mu = 1$. Suppose that holding cost $x$ is incurred per (discrete) unit time, when the number of customers in the system is $x$, and action 1 is used:

$$c_x(\phi^1_x) = c_x(1) = x, \quad x \in \mathbf{S}.$$

In [12] it was shown that the equation $v+g = c(\phi^1) + P(\phi^1)v$ has the following solutions: to any $g \in \mathbb{R}$ there is a solution tuple $(g, v_g)$ with $v^g : \mathbf{S} \to \mathbb{R}$ the function given by

$$\begin{aligned} v^g_x = &-\frac{x-1}{\mu-\lambda}g + \frac{(x-1)(x-2)}{2(\mu-\lambda)} + \mu\frac{x-1}{(\mu-\lambda)^2} + \frac{g}{\lambda} \\ &+ \mu\frac{(\mu/\lambda)^{x-1}-1}{\mu-\lambda}\Big\{\frac{g}{\mu-\lambda} + \frac{g}{\lambda} - \frac{\mu}{(\mu-\lambda)^2}\Big\}, \end{aligned} \tag{2.7}$$

for $x \geq 1$ and $v^g_0 = 0$. The solution obtained from a vanishing discount approach, is the one for which the expression between curly brackets is 0, i.e. for which

$$\frac{g}{\mu-\lambda} + \frac{g}{\lambda} - \frac{\mu}{(\mu-\lambda)^2} = 0,$$

in other words

$$g = \frac{\lambda}{\mu-\lambda},$$

and

$$v^g = -\frac{x-1}{\mu-\lambda}g + \frac{(x-1)(x-2)}{2(\mu-\lambda)} + \mu\frac{x-1}{(\mu-\lambda)^2} + \frac{g}{\lambda}.$$

This can also be derived from boundedness conditions analysed in [13]. Thus, $\mathsf{g}(\phi^1) = \lambda/(\mu-\lambda)$.

Next, in state 1 there is a further option to choose action 2, with corresponding transition probabilities

$$p_{1,3}(2) = \lambda = 1 - p_{1,0}(2).$$

This yields parameter $\phi^2$, with $\phi^2_1 = 2$, and $\phi^2_x = 1$, $x \neq 1$. The corresponding cost $c_1(2)$ is chosen small enough (possibly negative) so that $\mathtt{g}(\phi^2) < \mathtt{g}(\phi^1)$. This yields an MDP satisfying Assumption 2.2.3. Although the direct cost possibly is not non-negative, it is bounded below.

We claim that we may choose a constant $g$ in Eq. (2.7) so large that

$$v^g_1 + g = c_1(1) + \sum_y p_{1y}(1)v^g_y < c_1(2) + \sum_y p_{1y}(2)v^g_y = c_1(2) + \lambda v^g_3 + \mu v^g_0,$$

in other words, the minimisation prescribes to choose action 1 in state 1. Indeed, this choice is possible if

$$c_1(2) + \lambda v^g_3 > 1 + \lambda v^g_2,$$

or

$$1 - c_1(2) < \lambda(v^g_3 - v^g_2). \tag{2.8}$$

It can be checked that

$$v^g_3 - v^g_2 > \frac{\mu^2}{\lambda^2}\Big(\frac{g}{\lambda} - \frac{\mu}{(\mu-\lambda)^2}\Big).$$

Therefore, one may choose $g > \mathtt{g}(\phi^1)$ large enough for Eq. (2.8) to be true. Hence, $(g, v^g)$ is a solution to Eq. (2.6) for the MDP with minimising policy $\phi^1$. However, by construction $g > \mathtt{g}(\phi^1) > \mathtt{g}(\phi^2)$. Thus, $(g, v^g)$ is a solution to the DAOE, where $g$ is not the minimum expected average cost, and the policy choosing the minimising action is not the optimal policy.

## 2.3 Continuous time Model

In this section we will consider continuous time parametrised Markov processes. The setup is analogous to the discrete time case. Again we consider a parameter space $\Phi$ and a countable state space $\mathbf{S}$. With each $\phi \in \Phi$ we associate an $\mathbf{S} \times \mathbf{S}$ generator matrix or $q$-matrix $Q(\phi)$ and a cost rate vector $c(\phi) : \mathbf{S} \to \mathbb{R}$. Following the construction in [44], see also [51], one can define a measurable space $(\Omega, \mathcal{F})$, a stochastic process $X : \Omega \to \{f : [0,\infty) \to \mathbf{S} \mid f \text{ right-continuous}\}$, a filtration $\{\mathcal{F}_t\}_t \subset \mathcal{F}$ to which $X$ is adapted, and a probability distribution $\mathsf{P}_\nu^\phi$ on $(\Omega, \mathcal{F})$, such that $X$ is the minimal Markov process with $q$-matrix $Q(\phi)$, for each initial distribution $\nu$ on $\mathbf{S}$ and $\phi \in \Phi$. Denote by $P(\phi) = \{ p_{t,xy}(\phi)\}_{x,y\in\mathbf{S}}$, $t \geq 0$, the corresponding minimal transition function and by $\mathsf{E}_\nu^\phi$ the expectation operator corresponding to $\mathsf{P}_\nu^\phi$.

**Assumption 2.3.1.** **i)** $Q(\phi)$ is a conservative, stable $q$-matrix, i.e. for $x \in \mathbf{S}$ and $\phi \in \Phi$

- $0 \leq q_x(\phi) = -q_{xx}(\phi) < \infty$;
- $\sum_y q_{xy}(\phi) = 0$.

**ii)** $\{ P_t(\phi)\}_{t\geq 0}$ is standard, i.e. $\lim_{t\downarrow 0} p_{t,xy}(\phi) = \delta_{xy}$, with $\delta_{xy}$ the Kronecker delta.

**iii)** $\phi \mapsto q_{xy}(\phi)$ and $\phi \mapsto c_x(\phi)$ are continuous, $x, y \in \mathbf{S}$;

**iv)** $\Phi$ is locally compact.

Let $\Phi' \subset \Phi$. The definition of the product property of $\{Q(\phi)\}_\phi$ and $\{c(\phi)\}_\phi$ with respect to $\Phi'$ is completely analogous to Definition 2.2.1. This entails $\Phi'$ to be compact in the product topology. Again, for easy reference, we say that $\Phi'$ has the product property if $\{Q(\phi)\}_\phi$ and $\{c(\phi)\}_\phi$ both have the product property with respect to $\Phi'$. If $\Phi$ has the product property, then the parametrised Markov process is an MDP. Analogously to Remark 2.2.1, $\Phi$ may represent the collection of stationary policies or the stationary, deterministic ones.

Suppose furthermore, that a lump cost is charged, in addition to a cost rate incurred per unit time. Say at the moment of a jump $x$ to $y$ lump cost $d_{xy}(\phi)$ is charged, when the parameter is $\phi$. This can be modelled as a (marginal) cost rate $c_x(\phi) = \sum_{y\neq x} d_{xy}(\phi) q_{xy}(\phi)$.

Below we give the definitions of various performance measures and optimality criteria. Later on we will provide conditions under which these exist.

For $\alpha > 0$, under parameter $\phi \in \Phi$ the *expected total $\alpha$-discounted cost value function $v^\alpha$* is given by

$$v_x^\alpha(\phi) = \mathsf{E}_x^\phi\Big[\int_{t=0}^\infty e^{-\alpha t} c_{X_t} dt\Big], \quad x \in \mathbf{S}.$$

Suppose that $\Phi' \subset \Phi$ has the product property. The *minimum expected total $\alpha$-discounted cost* w.r.t $\Phi'$ is defined as

$$v_x^\alpha = \inf_{\phi\in\Phi'} \{v_x^\alpha(\phi)\}, \quad x \in \mathbf{S}.$$

If $v^\alpha(\phi) = v^\alpha$, then $\phi$ is said to be optimal in $\Phi'$.

The *expected average cost* under parameter $\phi$ is given by

$$\mathsf{g}_x(\phi) = \limsup_{T\to\infty} \frac{1}{T} \mathsf{E}_x^\phi \Big[ \int_{t=0}^T c_{X_t} dt \Big], \quad x \in \mathbf{S}.$$

Suppose that $\Phi' \subset \Phi$ has the product property. The *minimum expected average cost* is defined as

$$\mathsf{g}_x = \inf_{\phi\in\Phi'} \{\mathsf{g}_x(\phi)\}, \quad x \in \mathbf{S}.$$

If $\mathsf{g}(\phi) = \mathsf{g}$ for some $\phi \in \Phi'$ then $\phi$ is said to be average cost optimal in $\Phi'$.

The notions of *Blackwell optimality* and *strong Blackwell optimality* are defined completely analogously to the discrete time versions.

A well-known procedure to determine the structure of an optimal policy in the continuous time case, is to reduce the continuous time MDP to a discrete time MDP in order to be able to apply VI. There are different time-discretisation methods. One is to consider the embedded jump process. Sometimes this is a viable method, see [35] where this approach has been taken. In Section 2.3.4 we give an example where the embedded jump approach seems to be less amenable to apply.

Instead, one may use uniformisation. However, applicability hinges on models, where the jumps are bounded as a function of parameter and state:

$$q := \sup_{x\in\mathbf{S},\phi\in\Phi} q_x(\phi) < \infty. \tag{2.9}$$

This property is violated in models with reneging customers, population models etc, and we will consider how to handle this next.

Let us first recall the uniformisation procedure.

## 2.3.1 Uniformisation

A detailed account of the uniformisation procedure and proofs can be found in [64]. If a continuous time parametrised Markov process has bounded transition rates (cf. Eq. (2.9)), it admits a transformation to an equivalent discrete time parametrised Markov process. Below we list the transformations for the $\alpha$-discounted and average cost cases.

For the discounted cost criterion the equivalent discrete time process is given by

$$P(\phi) = I + \frac{1}{q} Q(\phi), \qquad c^d(\phi) = \frac{c(\phi)}{\alpha+q}, \qquad \alpha^d = \frac{\alpha}{\alpha+q}, \qquad \phi \in \Phi. \tag{2.10}$$

Denote the discrete time $\alpha^d$-discounted cost under policy $\phi$ as $v^{d,\alpha^d}(\phi)$. Both the discrete-time and continuous time processes have equal expected cost, i.e. $v^{d,\alpha^d}(\phi) = v^\alpha(\phi)$. If $\Phi' \subset \Phi$ has the product property, then this implies that the optimal $\alpha$- and $\alpha^d$-discounted value functions with respect to $\Phi'$ are equal:

$$v^{d,\alpha^d} = v^\alpha.$$

For the average cost criterion the equivalent discrete time process is given by

$$P(\phi) = I + \frac{1}{q}Q(\phi), \qquad c^d(\phi) = \frac{c(\phi)}{q}, \qquad \phi \in \Phi.$$

Denote the discrete time average cost under parameter $\phi$ as $\mathsf{g}^d(\phi)$ and the value function as $v^d(\phi)$. Under the same parameter, the discrete-time and continuous time expected cost, relate to each other as follows

$$q\mathsf{g}^d(\phi) = \mathsf{g}(\phi).$$

The corresponding value functions are identical:

$$v^d(\phi) = v(\phi).$$

These relations apply similarly to optimal parameters in a product set $\Phi' \subset \Phi$.

The main concern is how to proceed in the case of unbounded jump rates $q = \infty$, when the above procedure is not possible.

### 2.3.2 Discounted cost

First we treat the discounted cost criterion. This section summarises the results of Chapter 3. That chapter only treats optimality within the class of stationary Markov policies, as we do in the present chapter. We recall some definitions. These definitions are closely related to the conditions used in the discrete time analysis in Section 2.2.1.

**Definition 2.3.1.**
- The function $W : \mathbf{S} \to (0,\infty)$ is said to be a *moment function*, if there exists an increasing sequence $\{K_n\}_n \subset \mathbf{S}$ of finite sets with $\lim_n K_n = \mathbf{S}$, such that $\inf_{x \notin K_n} W_x \to \infty$, as $n \to \infty$.
- The function $V : \mathbf{S} \to (0,\infty)$ is called a $(\gamma,\Phi)$-drift function if $Q(\phi)V \leq \gamma V$ for all $\phi \in \Phi$, where $QV_x := \sum_{y \in \mathbf{S}} q_{xy} V_y$.

**Assumption 2.3.2 ($\boldsymbol{\alpha}$). i)** There exist a constant $\gamma < \alpha$ and function $V : \mathbf{S} \to (0,\infty)$ such that $V$ is a $(\gamma,\Phi)$-drift function;

**ii)** $\sup_\phi \|c(\phi)\|_V =: c_V < \infty$ for all $\phi \in \Phi$;

**iii)** There exist a constant $\theta$ and a function $W : \mathbf{S} \to (0, \infty)$ such that $W$ is a $(\theta, \Phi)$-drift function and $W/V$ is a moment function, where $(W/V)_x = W_x/V_x$, $x \in \mathbf{S}$.

Assumptions 2.3.2 ($\alpha$) (i) and 2.3.2 ($\alpha$)(ii) are the continuous time counterpart of Assumption 2.2.2 ($\alpha$). Assumption 2.3.2($\alpha$)(iii) is sufficient to guarantee nonexplosiveness of the parametrised Markov process (cf. [67, Theorem 2.1]), and implies continuity properties of the map $\phi \mapsto (P_t(\phi)V)_x$, $x \in \mathbf{S}$.

**Theorem 2.3.1** (Theorem 3.4.1)**.** *Suppose that Assumptions 2.3.1 and 2.3.2* ($\alpha$) *hold, then* $\phi \mapsto v^\alpha(\phi)$ *is component-wise continuous and* $v^\alpha(\phi)$ *is the unique solution in* $\ell^\infty(\mathbf{S}, V)$ *to*

$$\alpha u = c(\phi) + Q(\phi)u.$$

**Theorem 2.3.2** (Theorem 3.4.2)**.** *Assume that* $\Phi' \subset \Phi$ *has the product property. Suppose further that Assumptions 2.3.1, and 2.3.2*($\alpha$) *hold. Then* $v^\alpha$ *is the unique solution in* $\ell^\infty(\mathbf{S}, V)$ *to the* $\alpha$*-discount optimality equation (CDOE)*

$$\alpha u_x = \inf_{\phi_x \in \Phi'_x} \{c_x(a_x) + \sum_y q_{xy}(\phi_x)u_y\}, \quad x \in \mathbf{S}. \tag{2.11}$$

*There exists* $\phi^* \in \Phi'$ *with* $\phi^*_x \in \arg\min_{\phi_x \in \Phi'_x} \{c_x(a_x) + \sum_y q_{xy}(\phi_x)u_y\}$, $x \in \mathbf{S}$. *Any policy* $\phi' \in \Phi'$ *that minimises Eq. (2.11) is optimal in* $\Phi'$*, and it holds that* $v^\alpha(\phi') = v^\alpha$.

As discussed in Section 2.2.2, the parameter set may contain a perturbation component. Introducing a perturbation yields a parameter set of the following form $\Phi = \mathcal{N} \times \mathcal{D}$, where $N$ is a perturbation parameter and $\mathcal{D}$ the set of deterministic stationary (or merely stationary) policies.

**Corollary 2.3.3** (cf. Theorem 3.5.1)**.** *Let* $\Phi = \mathcal{N} \times \mathcal{D}$. *Suppose Assumptions 2.3.1 and 2.3.2*($\alpha$) *hold. Assume that* $\{N\} \times \mathcal{D}$ *has the product property for* $N \in \mathcal{N}$. *Then,*

**i)** $\lim_{N \to N_0} v^\alpha(N) = v^\alpha(N_0)$.

**ii)** *Any limit point of* $(\delta^*_N)_{N \to N_0}$ *is optimal in* $\{N_0\} \times \mathcal{D}$.

**iii)** *Suppose that the MDP with parameter set* $\{N\} \times \mathcal{D}$ *is uniformisable, i.e.*

$$q^N := \sup_{x \in \mathbf{S}, \delta \in \mathcal{D}} |q_{xx}(N, \delta)| < \infty.$$

*Consider the discount discrete-time uniformised MDP, with transition matrices, cost and discount factor given by (cf. Eq. (2.10))*

$$P(N,\delta) = I + \frac{1}{q^N}Q(N,\delta), \quad c(N,\delta) = \frac{c(N,\delta)}{\alpha + q^N}, \quad \alpha^d = \frac{\alpha}{\alpha + q^N}.$$

*Then the MDP satisfies Assumptions 2.2.1 and 2.2.2($\alpha^d$), for the same function $V$.*

*Proof.* Assertions i) and ii) are in fact Theorem 3.5.1, but they follow easily from Theorems 2.3.1 and 2.3.2. Assertion iii) is a direct verification. □

**Roadmap to structural properties** We finally have collected the tools to provide a scheme for the derivation of structural properties of an optimal policy and value function for a continuous time MDP with unbounded jump rates, provided the required conditions hold. Applications of this scheme are discussed in Chapters 3 and 4.

Let $\Phi' = \mathcal{D}$ be the set of stationary, deterministic policies, and $\Phi = \mathcal{N} \times \mathcal{D}$. Each set $\{N\} \times \mathcal{D}$ is assumed to have the product property, $N \in \mathcal{N}$.

**Roadmap for $\alpha$-discounted MDPs in continuous time**

1. If Assumptions 2.3.1 and 2.3.2($\alpha$) hold, and $q < \infty$, do
   - **a)** perform a uniformisation;
   - **b)** use VI Algorithm 1 to verify the structural properties of an optimal policy and value function;
   - **c)** use the equivalence of uniformised and non-uniformised systems to obtain the structure of an optimal policy and value function of the non-uniformised continuous time MDP.
2. If Assumptions 2.3.1 and 2.3.2($\alpha$) hold, and $q = \infty$, do
   - **i)** perform a bounded jump perturbation leaving the structural properties intact and satisfying Assumptions 2.3.1 and 2.3.2($\alpha$). For instance, one might apply SRT (see Section 2.2.2) or try a brute force perturbation;
   - **ii)** do steps a,b,c. This potentially yields structural properties of an optimal policy and the value function for each $N$-perturbed MDP;

**iii)** take the limit for the perturbation parameter to vanish. Corollary 2.3.3 gives the structural results for an optimal policy and value function.

3. If the assumptions do not hold, or if no structural properties can be derived, then the outcome is inconclusive.

As has been mentioned already, one might apply discounted VI directly to the associated discrete time MDP, embedded on the jumps of the continuous time MDP (cf. e.g. [35, Theorem 4.12]). In the example of Section 2.3.4 we discuss some problems with the application of this procedure.

### 2.3.3 Average cost

We finally turn to studying the average cost criterion in continuous time. The assumptions that we make, are Assumption 2.3.1 and the analog of Assumption 2.2.3 that we used in Section 2.2.3 for analysing the average cost criterion in discrete time. In fact, Assumption 2.2.3 can be used unaltered. However, one has to use the continuous time definitions of the hitting time of a state, and total expected cost incurred till the hitting time.

The hitting time $\tau_z$ of a state $z \in \mathbf{S}$ is defined by:

$$\tau_z = \inf_{t>0}\{X_t = z, \exists s \in (0,t) \text{ such that } X_s \neq z\}. \tag{2.12}$$

Then, $m_{xz}(\phi) = \mathsf{E}_x^\phi \tau_z$ and $c_{xz} = \mathsf{E}_x^\phi \int_0^{\tau_z} c_{X_t} dt$, where either expression may be infinite.

The following theorem is completely analogous to the discrete time equivalent, with the only difference that the CAOE below has a slightly different form.

**Theorem 2.3.4.** *Suppose, that Assumptions 2.3.1, 2.3.2($\alpha$), $\alpha > 0$, and 2.2.3 hold.*

**i)** *There exists a solution tuple $(g^*, v^*)$ to the average cost optimality equation (CAOE)*

$$g = \min_{\phi_x \in \Phi'_x} \{c_x(\phi_x) + \sum_{y \in \mathbf{S}} q_{xy}(\phi) u_y\}, \tag{2.13}$$

*with the property that (1) $g^* = \mathsf{g}$ is the minimum expected average cost (in $\Phi'$), (2) $\phi^* \in \Phi'$ with*

$$\phi_x^* \in \arg\min_{\phi_x \in \Phi'_x} \{c_x(\phi_x) + \sum_{y \in \mathbf{S}} q_{xy}(\phi_x) u_y\}$$

*is (average cost) optimal in $\Phi'$, and (3) there exists $x^* \in D$ with $v^*_{x^*} = \inf_x v^*_x$.*

**ii)** *Let $x_0 \in \mathbf{S}$. Any sequence $\{\alpha_n\}_n$ with $\lim_{n\to\infty} \alpha_n = 0$, has a subsequence, again denoted $\{\alpha_n\}_n$, along which the following limits exist:*

$$\begin{aligned} v'_x &= \lim_{n\to\infty} \{v^{\alpha_n}_x - v^{\alpha_n}_{x_0}\}, \quad x \in \mathbf{S}, \\ g' &= \lim_{n\to\infty} \alpha_n v^{\alpha_n}_x, \quad x \in \mathbf{S}, \\ \phi' &= \lim_{n\to\infty} \phi^{\alpha_n}. \end{aligned}$$

*Furthermore, the tuple $(g', v')$ is a solution to (2.13) with the properties (1), (2) and (3), so that $g' = \mathtt{g}$. Moreover, $\phi'$ takes minimising actions in (2.13) for $g = g'$ and $v = v'$.*

We have not encountered the above result in this form. However, the derivations are analogous to the discrete time variant, cf. [63, Chapter 7], and to the proofs in [35], where continuous time variants of Sennott's discrete time conditions have been assumed. In fact, Assumption 2.2.3 implies [35, Assumption 5.4]. Although one could piece together the proof of Theorem 2.3.4 from these references, we prefer to give it explicitly in Section 2.3.5.

For the verification of Assumption 2.2.3 (ii) and Assumption 2.2.3 (iv) one may use the following lemma, that is analogous to Lemma 2.2.6. The proof is similar to the proof of [72, Theorem 1].

**Lemma 2.3.5.** *Let $x_0 \in \mathbf{S}$ be given. Let $\phi \in \Phi'$. Suppose that there exist functions $f, h : \mathbf{S} \to [0, \infty)$ with*

**i)** *$f_x \geq \max\{1, c_x(\phi)\}$, $x \in \mathbf{S} \setminus \{x_0\}$;*

**ii)** *$f_x + \sum_{\substack{y: y \neq x_0 \\ \text{if } x \neq x_0}} q_{xy}(\phi) h_y \leq 0$, $x \in \mathbf{S}$.*

*Then $m_{xx_0}(\phi), c_{xx_0}(\phi) \leq h_x$, $x \in \mathbf{S}$.*

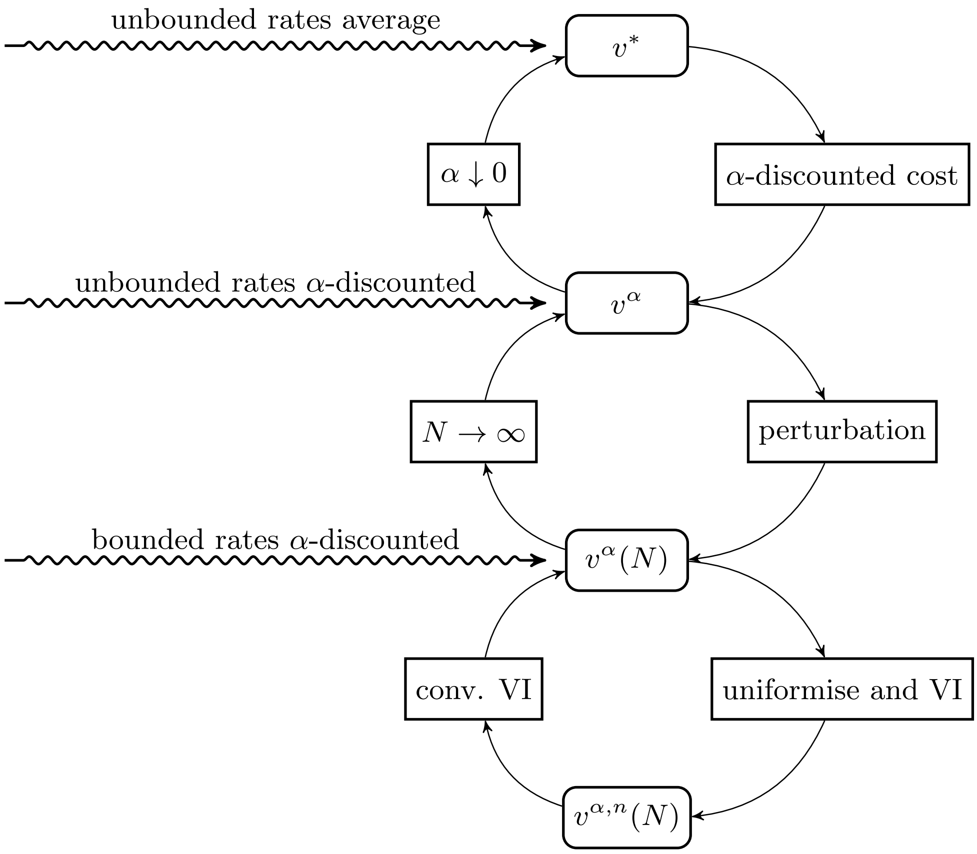
### 2.3.4 Roadmap to structural properties

First we present a roadmap for determining structural properties of average cost MDPs in continuous time. We illustrate it with a simple example. More complicated examples can be found in Chapters 4 and 5. Then a table will summarise the schematic approach that we have presented through the various roadmaps, including references to the required conditions and results.

Let $\Phi' = \mathcal{D}$ be the set of stationary, deterministic policies, and $\Phi = \mathcal{N} \times \mathcal{D}$. Assume that $\{N\} \times \mathcal{D}$ has the product property for $N \in \mathcal{N}$.

**Roadmap for average cost MDPs in continuous time**

1. If Assumptions 2.3.1, 2.3.2($\alpha$), for all $\alpha > 0$, and 2.2.3 hold then do
    - apply the roadmap for $\alpha$-discounted MDPs in continuous time; if the outcome is that the $\alpha$-discounted problem has the desired structural properties for all $0 < \alpha < \alpha_0$, for some $\alpha_0 > 0$, then do
    - apply the vanishing discount approach of Theorem 2.3.4 (ii).
2. If the assumptions do not hold, or structural properties can not be shown, the outcome is inconclusive.



The picture uses the following, hitherto not explained notation. If the perturbation parameter is $N$, then the $\alpha$-discounted value function is denoted by $v^{\alpha}(N)$. Applying discounted cost VI to the $N$-perturbation, yield the iterates $v^{\alpha,n}(N)$, for $n = 0, 1, \ldots$. The limit as the perturbation parameter vanishes is represented by $N \to \infty$.

**Arrival control of the M/M/1+M-queue** As an application of this final roadmap, we consider arrival control of the M/M/1+M-queue. Customers arrive in a single server unit with infinite buffer size according to a Poisson ($\lambda$) process. Each customer requires an exponentially distributed service time with parameter $\mu$, but he may also renege after an exponentially distributed amount of time with parameter $\beta$ (service is not exempted from reneging). Arrival process, service times and reneging times are all independent.

Due to reneging, the process associated with the number of customers in the server unit is ergodic at exponential rate. However, having reneging customers is not desirable from a customer service point of view. Therefore, the following arrival control is exercised. Per unit time and per customer a holding cost of size 1 is incurred. The controller can decide to accept (decision $\mathsf{A}$) or reject (decision $\mathsf{R}$) an arriving customer. If he takes decision $\mathsf{A}$, then a lump reward of size $K$ is incurred.

The goal is to select the control policy with minimum expected average cost. We wish to show that a control-limit acceptance policy is optimal. In other words, that there exists $x^* \in \mathbf{S}$, such that accepting in state $x \leq x^*$ and rejecting in state $x > x^*$ is average cost optimal.

This leads to the following MDP on the state space $\mathbf{S} = \mathbf{Z}_+$, where state $x$ corresponds to $x$ customers being present in the system. The collection of stationary, deterministic policies is given by $\mathcal{D} = \{\mathsf{A}, \mathsf{R}\}^{\infty}$. The transition rates are as follows: for $x \in \mathbf{S}$

$$q_{xy}(\mathsf{A}) = \begin{cases} \lambda, & y = x+1 \\ \mu \mathbb{1}_{\{x>0\}} + x\beta, & y = x-1 \\ -(\lambda + \mu \mathbb{1}_{\{x>0\}} + x\beta), & y = x, \end{cases}$$

$$q_{xy}(\mathsf{R}) = \begin{cases} \mu \mathbb{1}_{\{x>0\}} + x\beta, & y = x-1 \\ -(\mu \mathbb{1}_{\{x>0\}} + x\beta), & y = x. \end{cases}$$

The lump reward can be modelled as a cost rate, and we get for $x \in \mathbf{S}$

$$c_x(\mathsf{A}) = x - \lambda K, \qquad c_x(\mathsf{R}) = x.$$

This is an unbounded-rate MDP. Denote the never accepting policy by $\delta_0$, then this generates a Markov process with absorbing state 0, and finite expected average cost $\mathtt{g}(\delta_0) = 0$. One can check for $f_x = x$, $x \in \mathbf{S}$, that $h_x = e^{\theta x}$, $x \in \mathbf{S}$, satisfies Lemma 2.3.5 (ii), if we choose $\theta > \ln(1+\beta^{-1})$. Let $\epsilon > 0$. It then follows that Assumption 2.2.3 is satisfied with set $D = \{x \mid x - \lambda K \le 0 + \epsilon\}$.

It is not difficult to verify that Assumptions 2.3.1 and 2.3.2 $(\alpha)$, $\alpha > 0$, are satisfied. Indeed, for given $\alpha > 0$, there exists $\kappa_\alpha > 0$, such that $V^\alpha_x = e^{\kappa_\alpha x}$, $x \in \mathbf{S}$, is a $(\gamma_\alpha, \mathcal{D})$-drift function, for some $\gamma_\alpha > 0$.

It follows that there exists a solution tuple $(g^*, v^*)$ of the CAOE (2.13) with the properties (1), (2), (3). This CAOE takes the form

$$
\begin{aligned}
g^* = x + \big(\mu \mathbb{1}_{\{x>0\}} + (x \wedge N)\beta\big) v^*_{x-1} + \lambda \min\{-K + v^*_{x+1}, v^*_x\} \\
- \big(\lambda + \mu \mathbb{1}_{\{x>0\}} + (x \wedge N)\beta\big) v^*_x,
\end{aligned}
$$

where we have already rearranged the terms in such a way that the equation is amenable to analysis. It is easily deduced, that it is optimal to accept in state $x$ if

$$
v^*_{x+1} - v^*_x \le K.
$$

Hence, in order that a control-limit acceptance policy be average cost optimal, it is sufficient to show that $v^*$ is convex.

To this end, we will use the roadmap to show that there exists a solution pair $(g^*, v^*)$ to the CAOE (2.13) with properties (1), (2) and (3), and with $v^*$ a convex function. Theorem 2.3.4 justifies using the vanishing discount approach, and so it is sufficient to show convexity of the $\alpha$-discount value function $v^\alpha$, for all $\alpha > 0$ sufficiently small. Note that the imposed conditions for the roadmap for $\alpha$-discount MDPs are satisfied, since these are imposed as well for the assertions in Theorem 2.3.4, and these have been checked to hold.

The roadmap for the verification of structural properties of $v^\alpha$ prescribes to choose suitable perturbations. We consider a simple perturbation, where the reneging rates are truncated at $N\beta$ in states $x \ge N$, $N \ge 1$. The value $N = \infty$ then corresponds to the original MDP. Thus, $q_{xy}(N, \delta_x) = q_{Ny}(\delta_x)$, for $x \ge N$. A simple verification implies for $\Phi = \{1, 2, \ldots, \infty\} \times \mathcal{D}$, that $V^\alpha$ is a $(\gamma_\alpha, \Phi)$-drift function and that Assumptions 2.3.1 and 2.3.2 $(\alpha)$ are satisfied, $\alpha > 0$, for this extended parameter space.

Fix $\alpha > 0$ and $N \in \{1, 2, \ldots\}$. By virtue of Corollary 2.3.3 it is sufficient to check convexity of the $\alpha$-discount value function $v^\alpha(N)$, for the $N$-perturbation. Finally, by Theorem 2.2.3 it is sufficient to check convexity of $v^{\alpha,n}(N)$, which is the $n$-horizon approximation of $v^\alpha(N)$. Convexity of $v^{\alpha,n}(N)$ follows iteratively by putting $v^{\alpha,0}(N) \equiv 0$, and checking that convexity is propagated through the iteration step: for $x \in \mathbf{S}$

$$\begin{aligned} v_x^{\alpha,n+1}(N) =& x - \alpha\big(\mu \mathbb{1}_{\{x>0\}} + (x \wedge N)\beta\big) v_{x-1}^{\alpha,n}(N) \\ & + \alpha\lambda \min\{-K + v_{x+1}^{\alpha,n}(N), v_x^{\alpha,n}(N)\} \\ & + \alpha(1 - \lambda - \mu \mathbb{1}_{\{x>0\}} + (x \wedge N)\beta) v_x^{\alpha,n}(N). \end{aligned} \tag{2.14}$$

Event based dynamic programming (cf. [45] and Chapter 7) applied to Eq. (2.14) yields precisely the propagation of convexity.

**Associated embedded jump MDP** Instead of introducing a perturbation, we could have applied discounted VI to the associated $\alpha$-discounted embedded jump MDP. The assumptions that we have made, imply convergence to the $\alpha$-discounted value function (cf. [35, Theorem 4.14]). This yields the following VI-scheme:

$$\begin{aligned} \bar{v}_x^{\alpha,n+1} = \min\Big\{ & \frac{1}{\alpha + \mu \mathbb{1}_{\{x>0\}} + x\beta}\big(x + (\mu \mathbb{1}_{\{x>0\}} + x\beta)\bar{v}_{x-1}^{\alpha,n}\big), \\ & \frac{1}{\alpha + \lambda + \mu \mathbb{1}_{\{x>0\}} + x\beta}\big(x - \lambda K + \lambda \bar{v}_{x+1}^{\alpha,n} + (\mu \mathbb{1}_{\{x>0\}} + x\beta)\bar{v}_{x-1}^{\alpha,n}\big)\Big\}. \end{aligned}$$

First note that starting the iterations with the simple function $\bar{v}^{\alpha,0} \equiv 0$, only yields a convex function $\bar{v}^{\alpha,1}$,

$$\bar{v}_x^{\alpha,1} = \frac{x - \lambda K}{\alpha + \lambda + \mu \mathbb{1}_{\{x>0\}} + x\beta}, \quad x = 0, 1, \ldots,$$

under restrictions on the input parameters. In the minimisation one has to compare terms with different denominators. For showing convexity this is even more complicated, since one has to show that

$$\bar{v}_{x+2}^{\alpha,n+1} - \bar{v}_{x+1}^{\alpha,n+1} \geq \bar{v}_{x+1}^{\alpha,n+1} - \bar{v}_x^{\alpha,n+1},$$

given convexity of $\bar{v}^{\alpha,n}$, where each of these terms is a minimisation of two terms with different denominators. Already for this simple example it is not clear that this will work. Note that applying VI on the average cost embedded jump MDP has the same disadvantages. Additionally, one needs extra conditions (cf. Theorem 2.2.7) to ensure that average VI converges at all.

**Summary** The next table summarises the different roadmaps, with the appropriate references to the results justifying the various steps.

**Summarising table**

| Time | Criterion | Roadmap | | | |
|---|---|---|---|---|---|
| DT | disc. | VI1 | | | |
| | | Thm. 2.2.3 | | | |
| DT | average | VI2 | | | |
| Vgeo | | Thm. 2.2.7 | | | |
| DT | average | VDA | then VI1 | | |
| no Vgeo | | Thm. 2.2.5 | Thm. 2.2.3 | | |
| CT | disc. | UNI | then VI1 | | |
| bdd. | | § 2.3.1 | Thm. 2.2.3 | | |
| CT | disc. | PB | then UNI | then VI1 | |
| unb. | | Cor. 2.3.3 | § 2.3.1 | Thm. 2.2.3 | |
| CT | average | VDA | then UNI | then VI1 | |
| bdd. | | Thm. 2.3.4 | § 2.3.1 | Thm. 2.2.3 | |
| CT | average | VDA | then PB | then UNI | then VI1 |
| unb. | | Thm. 2.3.4 | Cor. 2.3.3 | § 2.3.1 | Thm. 2.2.3 |

| Abbreviations to summarising table | |
|---|---|
| Discrete time | DT |
| Continuous time | CT |
| Bounded or unbounded rates | bdd. or unb. |
| $\alpha$-discounted | disc. |
| Value iteration algorithm 1 or 2 | VI1 or VI2 |
| Vanishing discount approach | VDA |
| Uniformisation | UNI |
| Perturbation | PB |
| Conditions Theorem 2.2.7 | Vgeo |

## 2.3.5 Proofs

For the proof of Theorem 2.3.4 we will need a number of preparatory lemmas.

**Lemma 2.3.6.** *Suppose that Assumptions 2.3.1, 2.3.2*($\alpha$)*,* $\alpha > 0$*, and 2.2.3 hold. The following hold.*

**i)** *Let* $\mu(\phi_0)$ *denote the stationary distribution under parameter* $\phi_0$*, where* $\phi_0$ *has been specified in Assumption 2.2.3. Then* $\phi_0$ *has one closed class,* $R$

*say, that is positive recurrent. It holds, that*

$$\mathsf{g}(\phi_0) = \alpha \sum_R \mu_x(\phi_0) v_x^\alpha(\phi_0).$$

**ii)** *Let $\phi \in \Phi'$. Let $x \notin D$, and put $\tau := \tau_D$ to be the hitting time of $D$ (cf. Eq. (2.12)). Then*

$$v_x^\alpha(\phi) \geq \mathsf{E}_x^\phi\Big[\mathbb{1}_{\{\tau=\infty\}}\frac{\mathsf{g}(\phi_0)+\epsilon}{\alpha}+\mathbb{1}_{\{\tau<\infty\}}\big((1-e^{-\alpha\tau})\frac{\mathsf{g}(\phi_0)+\epsilon}{\alpha}+e^{-\alpha\tau}v_{X_\tau}^\alpha(\phi)\big)\Big]. \tag{2.15}$$

**iii)** *There exists $x_\alpha \in D$ with $v_{x_\alpha}^\alpha = \inf_x v_x^\alpha$.*

*Proof.* First we prove (i). By virtue of Assumption 2.2.3 (ii) the Markov process associated with $\phi_0$ has one closed class, which is positive recurrent. Furthermore, absorption into this class takes place in finite expected time and with finite expected cost, for any initial state $x \notin R$, since necessarily $x_0 \in R$.

Then we get

$$\begin{aligned}\sum_{x\in R}\mu_x(\phi_0)\mathsf{E}_x^{\phi_0}\big[c_{X_t}\big] &= \sum_{x\in R}\mu_x(\phi_0)\sum_{y\in R}p_{t,xy}(\phi_0)c_y(\phi_0)\\ &= \sum_{y\in R}c_y(\phi_0)\sum_{x\in R}\mu_x(\phi_0)\,p_{t,xy}(\phi_0)\\ &= \sum_{y\in R}c_y(\phi_0)\mu_y(\phi_0) = \mathsf{g}(\phi_0),\end{aligned}$$

where the interchange of summation is allowed by nonnegativity. This is used as well to justify the next derivation

$$\begin{aligned}\alpha\sum_{x\in R}\mu_x(\phi_0)v_x^\alpha(\phi_0) &= \alpha\sum_{x\in R}\mu_x(\phi_0)\mathsf{E}_x^{\phi_0}\big[\int_{t=0}^\infty e^{-\alpha t}c_{X_t}dt\big]\\ &= \alpha\int_{t=0}^\infty e^{-\alpha t}\sum_{x\in R}\mu_x(\phi_0)\mathsf{E}_x^{\phi_0}\big[c_{X_t}\big]dt\\ &= \alpha\int_{t=0}^\infty e^{-\alpha t}g(\phi_0)dt = \mathsf{g}(\phi_0).\end{aligned}$$

The proof of (ii) follows by splitting the $\alpha$-discounted cost into three terms, the first two of which represent the $\alpha$-discounted cost till $\tau$, in the respective cases $\tau = \infty$ and $\tau < \infty$, and the third is the cost starting from $\tau < \infty$:

$$\begin{aligned}
v_x^\alpha(\phi) =& \mathsf{E}_x^\phi\Big[\int_{t=0}^\infty e^{-\alpha t} c_{X_t}\,dt\Big] \\
\geq& \mathsf{E}_x^\phi\Big[\mathbb{1}_{\{\tau=\infty\}}\int_{t=0}^\infty e^{-\alpha t}dt(\mathsf{g}(\phi_0)+\epsilon) \\
&\qquad + \mathbb{1}_{\{\tau<\infty\}}\Big(\int_{t=0}^\tau e^{-\alpha t}dt(\mathsf{g}(\phi_0)+\epsilon)\int_{t=\tau}^\infty e^{-\alpha t}c_{X_t}dt\Big)\Big] \\
=& \mathsf{E}_x^\phi\Big[\mathbb{1}_{\{\tau=\infty\}}\frac{\mathsf{g}(\phi_0)+\epsilon}{\alpha} + \mathbb{1}_{\{\tau<\infty\}}\big((1-e^{-\alpha\tau})\frac{\mathsf{g}(\phi_0)+\epsilon}{\alpha} + e^{-\alpha\tau}v_{X_\tau}^\alpha(\phi)\big)\Big].
\end{aligned}$$

The inequality is due to the definitions of $D$ and $\tau$.

We finally prove (iii). Part (i) implies the existence of $z_\alpha \in R$ such that $\mathsf{g}(\phi_0) \geq \alpha v_{z_\alpha}^\alpha(\phi_0)$. Then there also exists a $y_\alpha \in D$ with $\mathsf{g}(\phi_0) \geq \alpha v_{y_\alpha}^\alpha(\phi_0)$. Indeed, suppose such $y_\alpha \in D$ does not exist. Then $v_y^\alpha(\phi_0) > \frac{\mathsf{g}(\phi_0)}{\alpha}$ for all $y \in D$. This leads to a contradiction, since by virtue of part (ii)

$$\frac{\mathsf{g}(\phi_0)}{\alpha} \geq v_{z_\alpha}^\alpha(\phi_0) \geq \mathsf{E}_{z_\alpha}^\phi\Big[(1-e^{-\alpha\tau})\frac{\mathsf{g}(\phi_0)+\epsilon}{\alpha} + e^{-\alpha\tau}v_{X_\tau}^\alpha(\phi_0)\Big] > \frac{\mathsf{g}(\phi_0)}{\alpha}.$$

Let $x_\alpha = \arg\min_{y\in D} v_y^\alpha$, and so $v_{x_\alpha}^\alpha \leq v_{x_\alpha}^\alpha(\phi_0) \leq \frac{g(\phi_0)}{\alpha}$. Then $x_\alpha = \arg\min_y v_y^\alpha$, because by Eq. (2.15) for any $x \notin D(\phi_0)$ and $\alpha$-discount optimal policy $\phi_\alpha$

$$\begin{aligned}
v_x^\alpha &= v_x^\alpha(\phi_\alpha) \\
&\geq \mathsf{E}_x^\phi\Big[\mathbb{1}_{\{\tau=\infty\}}\frac{\mathsf{g}(\phi_0)+\epsilon}{\alpha} + \mathbb{1}_{\{\tau<\infty\}}\big((1-e^{-\alpha\tau})\frac{\mathsf{g}(\phi_0)+\epsilon}{\alpha} + e^{-\alpha\tau}v_{X_\tau}^\alpha\big)\Big] \\
&\geq \mathsf{E}_x^\phi\Big[\mathbb{1}_{\{\tau=\infty\}}v_{x_\alpha}^\alpha + \mathbb{1}_{\{\tau<\infty\}}\big((1-e^{-\alpha\tau})v_{x_\alpha}^\alpha + e^{-\alpha\tau}v_{x_\alpha}^\alpha\big)\Big] \\
&= v_{x_\alpha}^\alpha.
\end{aligned}$$

□

**Lemma 2.3.7.** *Suppose that Assumptions 2.3.1, 2.3.2($\alpha$), $\alpha > 0$, and 2.2.3 hold. Let $\{\alpha_n\}_n$ be a positive sequence converging to 0. The following hold.*

**i)** *There exist a subsequence, call it $\{\alpha_n\}_n$ again, and $x_0 \in D$ such that $\alpha_n v_{x_0}^{\alpha_n} \leq \mathsf{g}(\phi_0)$, $n = 1, 2, \ldots$.*

**ii)** *There exist a constant $L$ and a function $M : \mathbf{S} \to (0,\infty)$, such that $-L \leq v_x^\alpha - v_z^\alpha \leq M_x$, $\alpha > 0$.*

*Proof.* To prove (i), note that Lemma 2.3.6 (iii) implies for all $n$ the existence of $x_{\alpha_n} \in D$, such that $v^{\alpha_n}_{x_{\alpha_n}} \le v^{\alpha_n}_x$, $x \in \mathbf{S}$. By Assumption 2.2.3 (iii) $D$ is finite, and so there exists $x_0 \in D$ and a subsequence of $\{\alpha_n\}_n$, that we may call $\{\alpha_n\}_n$ again, such that $x_{\alpha_n} = x_0$. Therefore by Lemma 2.3.6 (i), for all $n$

$$\alpha_n v^{\alpha_n}_{x_0} \le \alpha_n \sum_x \mu_x(\phi_0) v^{\alpha_n}_x \le \alpha_n \sum_x \mu_x(\phi_0) v^{\alpha_n}_x(\phi_0) = \mathsf{g}(\phi_0).$$

For the proof of (ii), take

$$M_x = c_{xz}(\phi_0), \quad L = \max_{y \in D} c_{zy}(\phi^y),$$

with $z$ and $\phi^y$ from Assumptions 2.2.3 (ii) and (iv). Let $\alpha > 0$. Let strategy $\phi$ follow $\phi_0$ until $z$ is reached, from then onwards it follows the $\alpha$-discount optimal policy $\phi_\alpha$. Then again by Assumption 2.2.3 (ii) we have

$$v^\alpha_x \le v^\alpha_x(\phi) \le c_{xz}(\phi_0) + v^\alpha_z(\phi_\alpha) = c_{xz}(\phi_0) + v^\alpha_z.$$

Notice that Assumptions 2.2.3 (iii) and (iv) yield $L < \infty$. According to Lemma 2.3.6 (iv) there is a minimum cost starting state $x_\alpha \in D$. Let $\phi'$ be the policy that uses policy $\phi^{x_\alpha}$ of Assumption 2.2.3 (iv) until hitting $x_\alpha$, after which $\phi'$ follows the $\alpha$-discount optimal policy $\phi_\alpha$. This yields,

$$v^\alpha_z - v^\alpha_x \le v^\alpha_z - \min_x v^\alpha_x \le v^\alpha_z(\phi') - v^\alpha_{x_\alpha} \le c_{zx_\alpha}(\phi^{x_\alpha}) \le L.$$

□

**Lemma 2.3.8.** *Suppose that Assumptions 2.3.1, 2.3.2($\alpha$), $\alpha > 0$, and 2.2.3 hold. Then,*

$$\limsup_{\alpha \downarrow 0} \alpha v^\alpha_x \le \mathsf{g}_x(\phi), \quad x \in \mathbf{S}, \phi \in \Phi'.$$

*Proof.* Let $\phi \in \Phi'$. We wish to apply Theorem 2.3.12 for $s(t) = \sum_y p_{t,xy}(\phi) c_y(\phi)$. First, Assumption 2.3.1, Assumption 2.3.2 ($\alpha$) and the dominated convergence theorem yield that $t \mapsto \sum_y p_{t,xy}(\phi) c_y(\phi)$ is continuous and $|v^\alpha_x(\phi)| < \infty$ (cf. Theorem 3.3.3). By Assumption 2.2.3 (i),

$$\sum_y p_{t,xy}(\phi) c_y(\phi), \quad v^\alpha_x(\phi) \ge 0, \quad x \in \mathbf{S}.$$

Then, $S(\alpha) = v^\alpha_x(\phi)$ and $\mathsf{g}_x(\phi) = \limsup_{T\to\infty} \frac{1}{T} S_T$. Hence Theorem 2.3.12 (1c) implies

$$\limsup_{\alpha \downarrow 0} \alpha v^\alpha_x(\phi) \le \mathsf{g}_x(\phi).$$

□

**Lemma 2.3.9** ([35, Theorem 5.2]). *Suppose that Assumptions 2.3.1, 2.3.2*($\alpha$), $\alpha > 0$, *and 2.2.3 hold. Let* $(g, v)$ *be a tuple, with* $g \in \mathbb{R}$ *and* $v : \mathbf{S} \to [-L, \infty)$, $x \in \mathbf{S}$, *and* $\phi \in \Phi'$ *be such that*

$$g \geq c_x(\phi) + \sum_y q_{xy}(\phi) v_y, \quad x \in \mathbf{S}.$$

*Then* $\mathsf{g}_x(\phi) \leq g$, $x \in \mathbf{S}$.

*Proof.* The proof is identical to the proof of [35, Theorem 5.2]. □

Now we have all results at hand to finish the proof of Theorem 2.3.4. The most important difficulty is to obtain the CAOE from a continuous time average cost optimality inequality (CAOI). To achieve this we have translated a very interesting argument used in [63, Chapter 7] for the discrete time case to continuous time.

*Proof of Theorem 2.3.4.* Let $\{\alpha_n\}_n > 0$ be a positive sequence converging to 0. Lemma 2.3.7(ii) implies that $-L \leq v_x^{\alpha_n} - v_z^{\alpha_n} \leq M_x$, for a constant $L$ and a function $M : \mathbf{S} \to (0, \infty)$, and $x \in \mathbf{S}$. Note that $[-L, M_x]$ is compact. By a diagonalisation argument, the sequence has a convergent subsequence, denoted $\{\alpha_n\}_n$ again, along which the limit exists for any $x \in \mathbf{S}$, say $v_x^{\alpha_n} - v_z^{\alpha_n} \to v'_x$, $x \in \mathbf{S}$.

Lemma 2.3.7(i) implies that there exists a further subsequence, again denoted $\{\alpha_n\}_n$, such that $0 \leq \alpha_n v_{x_0}^{\alpha_n} \leq \mathsf{g}(\phi_0)$, for some $x_0 \in D$. Compactness of $[0, \mathsf{g}(\phi_0)]$ implies existence of a limit point, say $g'$, along a subsequence, that in turn is denoted by $\{\alpha_n\}_n$.

By the above, $\alpha_n(v_y^{\alpha_n} - v_{x_0}^{\alpha_n}) \to 0$, and thus $\alpha_n v_y^{\alpha_n} \to g'$ for all $y \in \mathbf{S}$.

Since $\Phi'$ is compact, there is a final subsequence of $\{\alpha_n\}_n$, denoted likewise, such that $\{\phi^{\alpha_n}\}_n$, with $\phi^{\alpha_n}$ an $\alpha$-discount optimal policy, has a limit point $\phi'$ say. The tuple $(g', v')$ has property (3) from part (i) of the Theorem.

We will next show that this tuple is a solution to the following inequality:

$$g' \geq c_x(\phi') + \sum_y q_{xy}(\phi') v'_y \geq \inf_{\phi \in \Phi'} \{c_x(\phi) + \sum_y q_{xy}(\phi) v'_y\}. \tag{2.16}$$

Indeed, the $\alpha$-DDOE (2.11) yields for all $x \in \mathbf{S}$

$$\alpha v_x^\alpha \quad = \quad c_x(\phi_\alpha) + \sum_y q_{xy}(\phi_\alpha) v_y^\alpha.$$

Then we use Fatou's lemma and obtain

$$\begin{aligned}
g' &= \liminf_{n\to\infty}\{\alpha_n v_x^{\alpha_n}\} \\
&= \liminf_{n\to\infty}\Big\{c_x(\phi_{\alpha_n}) + \sum_{y\neq x} q_{xy}(\phi_{\alpha_n})[v_y^{\alpha_n} - v_z^{\alpha_n}] - q_x(\phi_{\alpha_n})[v_x^{\alpha_n} - v_z^{\alpha_n}]\Big\} \\
&\geq c_x(\phi') + \sum_{y\neq x} \liminf_{n\to\infty}\{q_{xy}(\phi_{\alpha_n})[v_y^{\alpha_n} - v_z^{\alpha_n}]\} - \liminf_{n\to\infty}\Big\{q_x(\phi_{\alpha_n})[v_x^{\alpha_n} - v_z^{\alpha_n}]\Big\} \\
&= c_x(\phi') + \sum_y q_{xy}(\phi')v'_y \\
&\geq \inf_{\phi\in\Phi'}\{c_x(\phi) + \sum_y q_{xy}(\phi)v'_y\},
\end{aligned}$$

where subtraction of $v_z^{\alpha_n}$ is allowed, since $Q(\phi)$ has row sums equal to zero. In the third equation we use continuity of $\phi \mapsto c_x(\phi)$ and $\phi \mapsto q_{xy}(\phi)$.

This allows to show that $(g', v')$ has property (1) from the Theorem and that $\phi'$ is optimal in $\Phi'$. Indeed, Lemma 2.3.8 and Lemma 2.3.9 yield for all $x \in \mathbf{S}$

$$\mathrm{g}_x(\phi') \leq g' = \lim_{n\to\infty} \alpha_n v_x^{\alpha_n} \leq \mathrm{g}_x \leq \mathrm{g}_x(\phi'). \tag{2.17}$$

Hence $\mathrm{g}_x(\phi') = \mathrm{g}_x = g'$, $x \in \mathbf{S}$, and so $\phi'$ is optimal in $\Phi'$, and $g'$ is the minimum expected average cost.

The following step is to show that both inequalities in Eq. (2.16) are in fact equalities. To this end, it is sufficient to show that $(g', v')$ is a solution tuple to the CAOE (2.13). Then Eq. (2.16) immediately implies that $\phi'$ takes minimising actions in Eq. (2.13) for the solution $(g', v')$.

Hence, let us assume the contrary. If $g' > \inf_{\phi\in\Phi'}\{c_x(\phi) + \sum_y q_{xy}(\phi)v'_y\}$ then there exists $\bar\phi_x \in \Phi'_x$, such that $g' > c_x(\bar\phi_x) + \sum_y q_{xy}(\bar\phi_x)v'_y$. Put $d_x \geq 0$ to be the corresponding discrepancy

$$g' = c_x(\bar\phi_x) + d_x + \sum_y q_{xy}(\bar\phi_x)v'_y.$$

As a consequence, if the inequality in Eq. (2.16) is not an equality, then there exists $\bar\phi \in \Phi'$ and a discrepancy function $d : \mathbf{S} \to [0, \infty)$, $d \not\equiv 0$, such that

$$g' = c_x(\bar\phi) + d_x + \sum_y q_{xy}(\bar\phi)v'_y, \quad x \in \mathbf{S}. \tag{2.18}$$

In other words

$$0 = c_x(\bar\phi) + d_x - g' + \sum_y q_{xy}(\bar\phi)v'_y, \quad x \in \mathbf{S}.$$

For $x \notin D$, $c_x(\bar{\phi}) + d_x - g' \geq \mathsf{g}(\phi_0) + \epsilon - g' \geq \epsilon$, and so $v' + L\mathbf{e}$ is a non-negative solution to the equation

$$\sum_y q_{xy}(\bar{\phi})(v'_y + L) \leq -\epsilon, \quad y \notin D.$$

This is precisely the condition in [72, Theorem 1] with $\lambda = 0$[2]. Following the proof of that theorem and using that $q_x(\bar{\phi}) > 0$ for $x \notin D$ (otherwise $\mathsf{g}_x(\bar{\phi}) = c_x(\bar{\phi}) > g'$), we can conclude that

$$v'_x + L \geq m_{xD}(\bar{\phi}), \quad x \notin D,$$

so that $m_{xD}(\bar{\phi}) < \infty$, for $x \notin D$.

For $x \in D$, either $q_x(\bar{\phi}) = 0$, or $q_x(\bar{\phi}) > 0$ and

$$m_{xD}(\bar{\phi}) = \frac{1}{q_x(\bar{\phi})} + \sum_{y \notin D} \frac{q_{xy}(\bar{\phi})}{q_x(\bar{\phi})} m_{yD}(\bar{\phi}) \leq \frac{1}{q_x(\bar{\phi})} + \sum_y \frac{q_{xy}(\bar{\phi})}{q_x(\bar{\phi})}(v'_y + L) < \infty,$$

by virtue of Eq. (2.18). We now will perform an iteration argument along the same lines as the proof of [72, Theorem 1].

First consider the case that $q_x(\bar{\phi}) > 0$. Dividing Eq. (2.18) for state $x$ by $q_x(\bar{\phi})$ we get, after reordering,

$$v'_x \geq \frac{c_x(\bar{\phi}) + d_x - g'}{q_x(\bar{\phi})} + \sum_{y \neq x} \frac{q_{xy}(\bar{\phi})}{q_x(\bar{\phi})} v'_y.$$

Introduce the substochastic matrix $P$ on $\mathbf{S} \setminus D$ by

$$p_{xy} = \begin{cases} \dfrac{q_{xy}(\bar{\phi})}{q_x(\bar{\phi})} & y \notin D \cup \{x\} \\ 0 & \text{otherwise.} \end{cases}$$

Denote the $n$ iterate by $P^{(n)}$, where $P^{(0)}$ is the $\mathbf{S} \times \mathbf{S}$ identity matrix. Then,

[2]The factor $\lambda$ in front of $y_i$ in that paper has been mistakenly omitted

for $x \notin D$ we get

$$\begin{aligned}
v'_x \geq & \frac{c_x(\bar{\phi}) + d_x - g'}{q_x(\bar{\phi})} + \sum_y p_{xy} v'_y + \sum_{y \in D} \frac{q_{xy}(\bar{\phi})}{q_x(\bar{\phi})} v'_y \\
\geq & \frac{c_x(\bar{\phi}) + d_x - g'}{q_x(\bar{\phi}} + \sum_{y \in D} \frac{q_{xy}(\bar{\phi})}{q_x(\bar{\phi})} v'_y \\
& + \sum_y p_{xy} \Big[ \frac{c_y(\bar{\phi}) + d_y - g'}{q_y(\bar{\phi})} + \sum_w p_{yw} v'_w + \sum_{w \in D} \frac{q_{yw}(\bar{\phi})}{q_y(\bar{\phi})} v'_w \Big] \\
\geq & \sum_{n=0}^{N-1} \sum_y p_{xy}^{(n)} \frac{c_y(\bar{\phi}) + d_y - g'}{q_y(\bar{\phi})} + \sum_{n=0}^{N-1} \sum_y p_{xy}^{(n)} \sum_{w \in D} \frac{q_{yw}(\bar{\phi})}{q_y(\bar{\phi})} v'_w + \sum_y p_{xy}^{(N)} v'_y .
\end{aligned}$$

Taking the liminf $N \to \infty$, we get

$$\begin{aligned}
v'_x \geq \sum_{n=0}^{\infty} \sum_y p_{xy}^{(n)} \frac{c_y(\bar{\phi}) + d_y - g'}{q_y(\bar{\phi})} + \sum_{n=0}^{\infty} \sum_y p_{xy}^{(n)} \sum_{w \in D} \frac{q_{yw}(\bar{\phi})}{q_y(\bar{\phi})} v'_w \\
+ \liminf_{N \to \infty} \sum_y p_{xy}^{(N)} v'_y .
\end{aligned}$$

Clearly

$$\liminf_{N \to \infty} \sum_y p_{xy}^{(N)} v'_y \geq \liminf_{N \to \infty} \sum_y p_{xy}^{(N)} (-L).$$

However, since $m_{xD}(\bar{\phi}) < \infty$, $x \notin D$, we get that $\liminf_{N \to \infty} \sum_y p_{xy}^{(N)} = 0$. Hence, for $\tau := \tau_D$

$$\begin{aligned}
v'_x &\geq \sum_{n=0}^{\infty} \sum_y p_{xy}^{(n)} \frac{c_y(\bar{\phi}) + d_y - g'}{q_y(\bar{\phi})} + \sum_{n=0}^{\infty} \sum_y p_{xy}^{(n)} xy \sum_{w \in D} \frac{q_{yw}(\bar{\phi})}{q_y(\bar{\phi})} v'_w \\
&\geq \mathsf{E}_x^{\bar{\phi}} \Big[ \int_{t=0}^{\tau} (c_{X_t} + d_{X_t} - g') dt \Big] + \mathsf{E}_x^{\bar{\phi}} \Big[ v'_{X_\tau} \Big] \\
&= c_{xD}(\bar{\phi}) + \mathsf{E}_x^{\bar{\phi}} \Big[ \int_{t=0}^{\tau} d_{X_t} dt \Big] - m_{xD}(\bar{\phi}) g' + \mathsf{E}_x^{\bar{\phi}} \Big[ v'_{X_\tau} \Big], \qquad (2.19)
\end{aligned}$$

for $x \notin D$. For $x \in D$ we can derive the same inequality. Note that we assumed $q_x(\bar{\phi}) > 0$. On the other hand, we have that

$$v_x^\alpha \leq c_{xD}(\bar{\phi}) + \mathsf{E}_x^{\bar{\phi}} \big[ e^{-\alpha\tau} v_{X_\tau}^\alpha \big].$$

This is equivalent to

$$v_x^\alpha - v_z^\alpha \le c_{xD}(\bar\phi) - v_z^\alpha(1 - \mathsf{E}_x^{\bar\phi}\big[e^{-\alpha\tau}\big]) + \mathsf{E}_x^{\bar\phi}\big[e^{-\alpha\tau}(v_{X_\tau}^\alpha - v_z^\alpha)\big].$$

Hence, for the sequence $\{\alpha_n\}_n$ we have

$$v_x^{\alpha_n} - v_z^{\alpha_n} \le c_{xD}(\bar\phi) - \alpha_n v_z^{\alpha_n}\frac{1 - \mathsf{E}_x^{\bar\phi}\big[e^{-\alpha_n\tau}\big]}{\alpha_n} + \mathsf{E}_x^{\bar\phi}\big[e^{-\alpha_n\tau}(v_{X_\tau}^{\alpha_n} - v_z^{\alpha_n})\big].$$

Taking the limit of $n$ to infinity yields

$$\begin{aligned} v_x' &\le c_{xD}(\bar\phi) - g'\cdot m_{xD}(\bar\phi) + \lim_{n\to\infty}\big\{\mathsf{E}_x^{\bar\phi}\big[e^{-\alpha_n\tau}(v_{X_\tau}^{\alpha_n} - v_z^{\alpha_n})\big]\big\} \qquad (2.20)\\ &= c_{xD}(\bar\phi) - g'\cdot m_{xD}(\bar\phi) + \mathsf{E}_x^{\bar\phi}\big[v'(X_\tau)\big]. \end{aligned}$$

Taking the limit through the expectation is justified by the dominated convergence theorem, since

$$|\mathsf{E}_x^{\bar\phi}\big[e^{-\alpha_n\tau}(v_{X_\tau}^{\alpha_n} - v_z^{\alpha_n})\big]| \le \mathsf{E}_x^{\bar\phi}e^{-\alpha_n\tau}|v_{X_\tau}^{\alpha_n} - v_z^{\alpha_n}| \le \mathsf{E}_x^{\bar\phi}(M_{X_\tau}\vee L) < \infty.$$

Combining Eqs. (2.19) and (2.20) yields $d \equiv 0$, for $x$ with $q_x(\bar\phi) > 0$.

For $x$ with $q_x(\bar\phi) = 0$, necessarily $g' = c_x(\bar\phi)$ and then equality in Eq. (2.16) immediately follows. Since there is equality for $\phi'$, this also implies that the inf is a min, and so we have obtained that $(g', v')$ is a solution to the CAOE (2.13).

The only thing left to prove, is that the solution tuple $(g', v')$ has property (2), that is, every minimising policy in Eq. (2.13) is average cost optimal. But this follows in the same manner as the argument leading to Eq. (2.17) yielding optimality of $\phi'$. This finishes the proof.

### 2.3.6 Tauberian Theorem

This section develops a Tauberian theorem that is used to provide the necessary ingredients for proving Theorem 2.3.4. This theorem is the continuous time counterpart of Theorem A.4.2 in Sennott [63]. A related assertion can be found in [35, Proposition A.5], however, in a weaker variant (without the Karamata implication, see Theorem 2.3.12, implication $(i) \implies (iii)$). The continuous time version seems deducible from Chapter 5 of the standard work on this topic [76]. We give a direct proof here.

Let $s : [0,\infty) \to \mathbb{R}$ be a function that is bounded below by $-L$ say and $(\mathcal{B}([0,\infty)),\mathcal{B})$-measurable, where $\mathcal{B}$ denotes the Borel-$\sigma$-algebra on $\mathbb{R}$, and $\mathcal{B}([0,\infty))$ the Borel-$\sigma$-algebra on $[0,\infty)$. Assume for any $\alpha > 0$ that

$$S(\alpha) = \int_{t=0}^{\infty} s(t)e^{-\alpha t}dt < \infty.$$

Furthermore, assume for any $T > 0$ that

$$S_T = \int_{t=0}^{T} s(t)dt < \infty.$$

**Lemma 2.3.10.** *Suppose that $L = 0$, i.e. $s$ is a nonnegative function. Then for all $\alpha > 0$ it holds that*

$$\frac{1}{\alpha}S(\alpha) = \int_{T=0}^{\infty} e^{-\alpha T} S_T dT. \tag{2.21}$$

*Furthermore, for all $\alpha > 0, U \geq 0$ the following inequalities hold true:*

$$\alpha S(\alpha) \geq \inf_{T \geq U} \left\{ \frac{S_T}{T} \right\} \left( 1 - \alpha^2 \int_{T=0}^{U} T e^{-\alpha T} dT \right), \tag{2.22}$$

*and*

$$\alpha S(\alpha) \leq \alpha^2 \int_{T=0}^{U} e^{-\alpha T} S_T dT + \sup_{T \geq U} \left\{ \frac{S_T}{T} \right\}. \tag{2.23}$$

*Proof.* We first prove Eq. (2.21). To this end, let $\alpha > 0$. Then,

$$\begin{aligned}
\frac{1}{\alpha}S(\alpha) &= \int_{u=0}^{\infty} e^{-\alpha u} du \int_{t=0}^{\infty} s(t) e^{-\alpha t} dt \\
&= \int_{t=0}^{\infty} \int_{u=0}^{\infty} s(t) e^{-\alpha(u+t)} du dt \\
&= \int_{T=0}^{\infty} \int_{t=0}^{T} s(t) e^{-\alpha T} du dT \\
&= \int_{T=0}^{\infty} e^{-\alpha T} \int_{t=0}^{T} s(t) du dT \\
&= \int_{T=0}^{\infty} e^{-\alpha T} S_T dT.
\end{aligned}$$

Interchange of integrals, change of variables are allowed, since the integrands are non-negative and the integrals are finite.

Next, we prove Eq. (2.22). To this end, we use Eq. (2.21). Then, for all

$\alpha > 0$, $U \geq 0$

$$\begin{aligned}\alpha S(\alpha) &= \alpha^2 \int_{T=0}^{\infty} S_T e^{-\alpha T} dT \\ &\geq \alpha^2 \int_{T=U}^{\infty} \frac{S_T}{T} T e^{-\alpha T} dT \\ &\geq \alpha^2 \inf_{t \geq U} \left\{ \frac{S_T}{T} \right\} \int_{T=U}^{\infty} T e^{-\alpha T} dT \\ &= \alpha^2 \inf_{t \geq U} \left\{ \frac{S_T}{T} \right\} \left( \int_{T=0}^{\infty} T e^{-\alpha T} dT - \int_{T=0}^{U} T e^{-\alpha T} dT \right) \\ &= \inf_{T \geq U} \left\{ \frac{S_T}{T} \right\} \left( 1 - \alpha^2 \int_{T=0}^{U} T e^{-\alpha T} dT \right).\end{aligned}$$

The first inequality uses explicitly that the integrand is non-negative.

Similarly we expand from Eq. (2.21) to get Inequality (2.23) as follows. Let $\alpha > 0, U \geq 0$. Then,

$$\begin{aligned}\alpha S(\alpha) &= \alpha^2 \int_{T=0}^{\infty} e^{-\alpha T} S_T dT \\ &= \alpha^2 \int_{T=0}^{U} e^{-\alpha T} S_T dT + \alpha^2 \int_{T=U}^{\infty} e^{-\alpha T} T \frac{S_T}{T} dT \\ &\leq \alpha^2 \int_{T=0}^{U} e^{-\alpha T} S_T dT + \sup_{T \geq U} \left\{ \frac{S_T}{T} \right\} \alpha^2 \int_{T=U}^{\infty} T e^{-\alpha T} dT \\ &\leq \alpha^2 \int_{T=0}^{U} e^{-\alpha T} S_T dT + \sup_{T \geq U} \left\{ \frac{S_T}{T} \right\} \alpha^2 \int_{T=0}^{\infty} T e^{-\alpha T} dT \\ &= \alpha^2 \int_{T=0}^{U} e^{-\alpha T} S_T dT + \sup_{T \geq U} \left\{ \frac{S_T}{T} \right\}.\end{aligned}$$

□

Let $f : [0, 1] \to \mathbb{R}$ be an integrable function, and define

$$S_f(\alpha) = \int_{t=0}^{\infty} e^{-\alpha t} f(e^{-\alpha t}) s(t) dt.$$

**Lemma 2.3.11.** *Assume that $L = 0$ and*

$$W := \liminf_{\alpha \downarrow 0} \alpha S(\alpha) = \limsup_{\alpha \downarrow 0} \alpha S(\alpha) < \infty.$$

*Let $r : [0,1] \to \mathbb{R}$ be given by*

$$r(x) = \begin{cases} 0 & x < 1/e \\ 1/x & x \geq 1/e. \end{cases}$$

*Then*

$$\lim_{\alpha \downarrow 0} \alpha S_r(\alpha) = \left( \int_{x=0}^{1} r(x)dx \right) \lim_{\alpha \downarrow 0} \alpha S(\alpha). \tag{2.24}$$

*Proof.* We first prove Eq. (2.24) for polynomial functions, then for continuous functions and finally for $r$. To show that Eq. (2.24) holds for polynomials, it is sufficient to prove it for $p(x) = x^k$. Thus,

$$\begin{aligned} \alpha S_p(\alpha) &= \alpha \int_{t=0}^{\infty} e^{-\alpha t} (e^{-\alpha t})^k s(t)dt \\ &= \frac{1}{k+1} \left[ \alpha(k+1) \int_{t=0}^{\infty} e^{-\alpha(k+1)t} s(t)dt \right] \\ &= \int_{x=0}^{1} p(x)dx \left[ \alpha(k+1) S(\alpha(k+1)) \right]. \end{aligned}$$

Taking the limit of $\alpha \downarrow 0$ proves Eq. (2.24) for polynomials. This is allowed because $W$ is finite. Next we show Eq. (2.24) for continuous functions. The Weierstrass approximation theorem (see [71, Section 13.33], [4]) yields that a continuous function $q$ on a closed interval can be arbitrary closely approximated by polynomials. Let $p$ such that $p(x) - \epsilon \leq q(x) \leq p(x) + \epsilon$ for $0 \leq x \leq 1$. Then,

$$\int_{x=0}^{1} p(x)dx - \epsilon \leq \int_{x=0}^{1} q(x)dx \leq \int_{x=0}^{1} p(x)dx + \epsilon.$$

$$\begin{aligned} S_{p-\epsilon}(\alpha) &= \int_{t=0}^{\infty} e^{-\alpha t} (p(e^{-\alpha t}) - \epsilon) s(t)dt \\ &= \int_{t=0}^{\infty} e^{-\alpha t} p(e^{-\alpha t}) s(t)dt - \epsilon \int_{t=0}^{\infty} e^{-\alpha t} s(t)dt \\ &= S_p(\alpha) - \epsilon S(\alpha). \end{aligned}$$

This implies

$$0 \leq S_{p+\epsilon}(\alpha) - S_{p-\epsilon}(\alpha) \leq 2\epsilon S(\alpha),$$

As $\epsilon$ approaches 0, finiteness of $W$ yields the result for continuous functions. In a similar manner $r$ can be approximated by continuous functions $q, q'$ with

$q' \le r \le q$ as follows

$$q(x) = \begin{cases} 0 & x < \frac{1}{e} - \delta \\ \frac{e}{\delta}x + e - \frac{1}{\delta} + & \frac{1}{e} - \delta \le x < \frac{1}{e} \\ \frac{1}{x} & x \ge \frac{1}{e}, \end{cases}$$

$$q'(x) = \begin{cases} 0 & x < \frac{1}{e} \\ \frac{e}{\gamma+\gamma^2 e}x + \frac{1}{\gamma+\gamma^2 e} & \frac{1}{e} \ge x > \frac{1}{e} + \gamma \\ 1/x & \frac{1}{e} + \gamma \ge x. \end{cases}$$

This proves Eq. (2.24). □

**Theorem 2.3.12.** *The following assertions hold.*

1. $\liminf_{T\to\infty} \frac{S_T}{T} \overset{(a)}{\le} \liminf_{\alpha\downarrow 0} \alpha S(\alpha) \overset{(b)}{\le} \limsup_{\alpha\downarrow 0} \alpha S(\alpha) \overset{(c)}{\le} \limsup_{T\to\infty} \frac{S_T}{T}$;
2. *the following are equivalent*

$$\begin{array}{rll} i) & \liminf_{\alpha\downarrow 0} \alpha S(\alpha) & = \limsup\limits_{\alpha\downarrow 0} \alpha S(\alpha) < \infty; \\ ii) & \liminf_{T\to\infty} \frac{S_T}{T} & = \limsup\limits_{T\to\infty} \dfrac{S_T}{T} < \infty; \\ iii) & \lim_{\alpha\downarrow 0} \alpha S(\alpha) & = \lim\limits_{T\to\infty} \dfrac{S_T}{T} < \infty. \end{array}$$

*Proof.* This proof is based on Sennott [63]. Clearly inequality ($b$) holds. So this leaves to prove inequalities ($a$) and ($c$).

Proof of inequality ($a$). First notice, that if we take $s \equiv M$ a constant function, then

$$\liminf_{T\to\infty} \frac{S_T}{T} = \liminf_{\alpha\downarrow 0} \alpha S(\alpha) = \limsup_{\alpha\downarrow 0} \alpha S(\alpha) = \limsup_{T\to\infty} \frac{S_T}{T}.$$

Therefore adding a constant $M$ to the function $s$ does not influence the result. Hence, it is sufficient to prove the theorem for nonnegative functions $s$. This means that the assumptions of Lemma 2.3.10 hold and we may use Inequality (2.22). Thus,

$$\inf_{T\ge U} \left\{ \frac{S_T}{T} \right\} \left( 1 - \alpha^2 \int_{T=0}^{U} T e^{-\alpha T} dT \right) \le \alpha S(\alpha).$$

Notice that $\lim_{\alpha\downarrow 0}\alpha^2\int_{T=0}^{U} Te^{-\alpha T}dT = 0$, hence taking the $\liminf$ as $\alpha \downarrow 0$ gives

$$\inf_{T\geq U}\frac{S_T}{T} \leq \liminf_{\alpha\downarrow 0}\alpha S(\alpha).$$

Now taking the limit $U \to \infty$ on both sides gives

$$\liminf_{T\to\infty}\frac{S_T}{T} \leq \liminf_{\alpha\downarrow 0}\alpha S(\alpha),$$

which yields the result. Using Inequality (2.23) of Lemma 2.3.10 and applying the same reasoning proves inequality $(c)$.

Next we prove part 2. Part 1 implies that $i) \Longleftarrow ii) \Longleftrightarrow iii)$. So it is sufficient to prove that $i) \Longrightarrow iii)$.
Assume that $i)$ holds, then we may invoke Lemma 2.3.11. First notice that $\int_{x=0}^{1} r(x)dx = 1$, hence Eq. (2.24) reduces to

$$\lim_{\alpha\downarrow 0}\alpha S_r(\alpha) = \lim_{\alpha\downarrow 0}\alpha S(\alpha).$$

Moreover,

$$\alpha S_r(\alpha) = \alpha\int_{t=0}^{\infty} e^{-\alpha t}s(t)e^{\alpha t}\mathbb{1}_{\{e^{-\alpha t}\geq e^{-1}\}}dt = \alpha\int_{t=0}^{1/\alpha} s(t)dt = \alpha S_{1/\alpha}$$

To complete the proof, we have

$$\lim_{\alpha\downarrow 0}\alpha S(\alpha) = \lim_{\alpha\downarrow 0}\alpha S_r(\alpha) = \lim_{\alpha\downarrow 0}\alpha S_{1/\alpha} = \lim_{T\to\infty}\frac{S_T}{T}.$$

□

**Acknowledgements** We would like to thank Sandjai Bhulai for introducing us to the illustrative tandem queue model in Section 2.2.2. Moreover he provided us with numerical results for Figures 2.2 and 2.3.

# 3 Parametrised Markov processes with discounted cost

This chapter is based on Blok and Spieksma [18], published.

## 3.1 Introduction

In this chapter we study convergence and continuity properties of a collection of parametrised continuous time Markov processes in countable state space with a discounted cost criterion. The parameter may represent a stationary or deterministic policy in a Markov decision process (MDP). It may also represent a perturbation of a Markov process. Or it can be a combination of both; i.e. control in a perturbed MDP.

The motivation for this chapter is our interest in MDPs with unbounded transition rates. In order to study structural properties the MDP has to be uniformisable. Structural properties of optimal policies and the value function follow from the propagation of these properties through a value iteration step. Note that often value iteration is applicable to the associated jump MDP. However, it is not clear that the desired structural properties propagate through the value iteration step in this case, since the expected sojourn times in the states may not be equal and so they may affect the resulting immediate costs and transition probabilities in an undesirable manner.

Hence, we wish to perturb the MDP in such a manner that it allows uniformisation and the structural properties are preserved. Therefore continuity in the parameter is necessary to infer properties of the original MDP from properties of the perturbed MDPs.

The conditions we impose on the Markov processes boil down to the existence of a transformation of the process, such that the transformed process is nonexplosive and moreover has a bounded cost function. These conditions should hold uniformly in the parameter and are expressed as drift conditions for the original Markov process as well as for the transformed process. Nonexplosiveness of the transformation guarantees continuity of the relevant performance measures as a function of the parameter, provided some standard continuity conditions hold.

The typical performance measure we have in mind is the discounted value function. If the parameter space has a product property the parametrised process is an MDP. The continuity of the value function implies the existence of a solution of the continuous time discount optimality equation (CDOE). We show that the solution provides a deterministic stationary optimal policy in the class of stationary policies. We do not study history dependent policies.

As an illustration we apply our results to the server farm with unbounded rates studied in [1]. In that paper it was shown that for bounded jump perturbations of the model a switching curve policy is optimal. However the unbounded jump case remained open, since till recently no theory was available to justify taking the limit of the perturbation parameter going to 0 – and the jumps becoming unbounded. In the present chapter we take the parameter space to be the product of the perturbation and control parameters. The obtained continuity results allow us to take the limit and show that a switching curve policy with the same structure is optimal.

The drift conditions that are used to show the existence of a solution of the CDOE, are related to the conditions used in [34], [36], [37], [38], [51] and [53]. These papers do not study convergence results, to the best of the authors' knowledge the only paper where convergence of perturbed MDPs is studied, is [52]. We want to emphasise that our aim has been to give minimal conditions for the drift criteria. In the one-parameter case our drift conditions are proven to be necessary (cf. [68]). Furthermore, we have tried to highlight the role that the various conditions play in the derivations. The conditions we impose are weaker than those used in the above mentioned papers. More detailed comparisons with the other drift conditions are given later in the chapter, in Section 3.4 and Remark 3.5.2.

The chapter is organised as follows. Section 3.2 introduces a so-called $V$-transformation and gives a characterisation of nonexplosiveness in terms of drift conditions. In Section 3.3 we develop conditions implying the continuity properties we will need. In Section 3.4 the two main theorems regarding the solution to the Poisson and optimality equation are stated. In Section 3.5 the translation to MDPs and perturbed MDPs is made. We provide an outline of the approach to get results for unbounded MDPs. Finally, in Section 3.6 we demonstrate this approach on the server farm model studied by [1]. See also Chapter 4 for a more extensive treatment of this model.

## 3.2 Basic settings

We will restrict our investigations to the following class of parametrised processes.

**Assumption 3.2.1.** For each $a \in \mathsf{A}$, $X(a)$ is a minimal, standard, stable Markov process, with right-continuous sample paths (with respect to the discrete topology), and with conservative $q$-matrix $Q(a) = (q_{xy}(a))_{x,y\in\mathbf{S}}$, i.e. for all $x \in \mathbf{S}, a \in \mathsf{A}$

1. $0 \le q_x(a) = -q_{x,x}(a) < \infty$;
2. $\sum_y q_{xy}(a) = 0$.

With $P_t(a) = \{\, p_{t,xy}(a)\}_{xy}$, $t \ge 0$, we denote the minimal transition function. A basic role in the discussion of relevant continuity properties of a parametrised Markov process is played by explosiveness properties. To this end we will first review the definition of explosiveness and a characterisation that is useful in this context. For the rest of this section we restrict to the one-parameter case.

We will define this properly. To this end, let $X$ be a Markov process on $\mathbf{S}$ that satisfies Assumption 3.2.1 (for a parameter space consisting of one element). Let $\tau_0 = 0$ and $\tau_{n+1} = \inf\{t > \tau_n \,|\, X_t \ne X_{t-}\}$ if $X_{\tau_n}$ is not-absorbing. Otherwise, put $\tau_k = \infty$ and $X_{\tau_k} = X_{\tau_n}$ for $k > n$. Let $J_\infty = \lim_{n\to\infty} \tau_k$. $X$ is said to be explosive if there exists a state $x \in \mathbf{S}$ with $\mathsf{P}\{J_\infty < \infty \,|\, X_0 = x\} > 0$. Nonexplosiveness is strongly related to the existence of a drift moment function, introduced below. First we need some notation. Let $f : \mathbf{S} \to \mathbb{R}$, then $f$ can be viewed as a vector of dimension $|\mathbf{S}|$. By $Qf$ and $P_t f$ we mean the matrix times vector products with elements $Qf(x) = \sum_{y\in\mathbf{S}} q_{xy} f(y)$, and $P_t f(x) = \sum_{y\in\mathbf{S}} p_{t,xy} f(y)$, $x \in \mathbf{S}$ respectively.

**Definition 3.2.1.** Let $\gamma \in \mathbb{R}$ and $V : \mathbf{S} \to \mathbb{R}_+ = (0,\infty)$, then

- $V$ is said to be a $\gamma$-drift function for $X$ if $QV \le \gamma V$, where we use component-wise ordering;
- $V$ is said to be a moment function, if there exists an increasing sequence $\{K_n\}_n \subset \mathbf{S}$ of finite sets with $\lim_n K_n = \mathbf{S}$, such that $\inf_{x\notin K_n} V(x) \to \infty$, as $n \to \infty$.

Note that since $Q$ is conservative, $V \equiv 1$ is always a 0-drift function. Furthermore, the paper [67, Theorem 2.1] shows that nonexplosiveness of $X$ is equivalent to the existence of a $\gamma$-drift moment function, for some constant $\gamma \in \mathbb{R}$.

**Definition 3.2.2.** Let $\gamma \in \mathbb{R}$, $V$ be a $\gamma$-drift function for $X$. Define the following associated transformation of $X$, denoted as $X^V$. Extend the state space with a coffin state $\Delta$, i.e. $\mathbf{S}_\Delta = \mathbf{S} \cup \{\Delta\}$. Then define

$$q^V_{xy} = \begin{cases} \frac{q_{xy}V(y)}{V(x)}, & x \neq y, x, y \neq \Delta \\ q_{xx} - \gamma, & x = y, x, y \neq \Delta \\ \gamma - \frac{\sum_{y \in \mathbf{S}} q_{xy}V(y)}{V(x)}, & x \neq \Delta, y = \Delta \\ 0, & x = \Delta, y \in \mathbf{S}_\Delta. \end{cases}$$

This makes $Q^V = (q^V_{xy})_{x,y \in \mathbf{S}_\Delta}$ a conservative $q$-matrix, with $\Delta$ an absorbing state. Denote by $\{P^V_t\}_t$ again the (minimum) transition function on the enlarged state space $\mathbf{S}_\Delta$.

Since we also need to take into account a cost or reward structure, the validity of the Kolmogorov forward integral equation is an important tool in guaranteeing the existence of solutions to CDOEs. The function $f : \mathbf{S} \to \mathbb{R}$ is said to satisfy the Kolmogorov forward equation if for all $x \in \mathbf{S}$

$$P_t f(x) = f(x) + \int_0^t P_s(Qf)(x) ds, \quad t \geq 0, \tag{3.1}$$

where $P_t f(x) = \sum_y p_{t,xy} f(y)$.

The following result holds.

**Theorem 3.2.1** (cf. [66, Theorem 3.2], [67, Theorem 2.1])**.** *Let Assumption 3.2.1 hold and let $V$ be a $\gamma$-drift function for $X$. The following are equivalent*

**i)** *$V$ satisfies Eq. (3.1);*

**ii)** *$X^V$ is nonexplosive;*

**iii)** *for some constant $\theta$ there exists a $\theta$-drift $V$-moment function $W$ for $X$.*

With $W$ being a $V$-moment function we mean that $W/V$ is a moment function. Then direct calculations yield that $W$ being a $\theta$-drift $V$-moment function for $X$ is equivalent to $W/V$ being a $(\theta-\gamma)$-drift moment function for $X^V$, where $(W/V)(x) = W(x)/V(x)$, $x \in \mathbf{S}$.

Under any of these three conditions, the functions bounded by $V$ also satisfy Eq. (3.1), under suitable integrability conditions. A discounted version is needed later on, and so we make it precise in the theorem below. To do so, we need some further notation.

The Banach space of functions bounded by $V$ (or $V$-bounded functions) on $\mathbf{S}$ is denoted by $\ell^\infty(\mathbf{S}, V)$. This means that $f \in \ell^\infty(\mathbf{S}, V)$ if $f : \mathbf{S} \to \mathbb{R}$ and

$$\|f\|_V := \sup_{x \in \mathbf{S}} \frac{|f(x)|}{V(x)} < \infty.$$

If $V$ is a $\gamma$-drift function, then [3] implies $P_t V \le e^{\gamma t} V$ and [67] implies $t \mapsto P_t V$ is continuous on $\mathbb{R}_+$. This implies that $t \mapsto P_t f$ is continuous for each $f \in \ell^\infty(\mathbf{S}, V)$ and, hence, integrable. Additionally, $P_t$ is a $V$-bounded linear operator, mapping $\ell^\infty(\mathbf{S}, V)$ into itself, with induced norm

$$\| P_t\|_V = \sup_x \frac{P_t V(x)}{V(x)} \le e^{\gamma t}. \tag{3.2}$$

Note that in general the $q$-matrix $Q$ is not a $V$-bounded linear operator.

**Theorem 3.2.2** (cf. [66, Theorem 3.4, Lemma 3.1]). *Let Assumption 3.2.1 hold, and let $V$ be a $\gamma$-drift function for $X$.*

**i)** *If $X^V$ is nonexplosive and, either $f \in \ell^\infty(\mathbf{S}, V)$ and $\int_0^t P_s|Qf|ds < \infty$ or $f = V$, then for any $k \in \mathbb{R}$, $f$ satisfies*

$$e^{kt} P_t f(x) = f(x) + \int_0^t e^{ks} \Big[ P_s(Qf)(x) + k\, P_s f(x)\Big] ds. \tag{3.3}$$

**ii)** *Conversely, if $V$ satisfies Eq. (3.3) for some $k \in \mathbb{R}$, then $X^V$ is nonexplosive.*

*Proof.* The proof of Theorem 3.2.2(i) follows entirely from the proofs in the referenced theorem and lemma. The conditions in the referenced results are slightly different: $f$ is assumed to be a $\gamma'$-drift function for some $\gamma' \in \mathbb{R}$. However, this is only used in the proofs to guarantee that $\int_0^t P_s|Qf|(x)ds < \infty$. The latter is assumed explicitly here.

For the proof of Theorem 3.2.2(ii), we assume that Eq. (3.3) holds for $V$. By virtue of [3, Lemma 5.4.2] we have

$$p_{t,xy} = \frac{V(x)}{V(y)} e^{\gamma t}\, p^V_{t,xy}, \quad x, y \in \mathbf{S}, t \ge 0. \tag{3.4}$$

Hence, we can rewrite Eq. (3.3) as

$$e^{kt} \sum_{y \in \mathbf{S}} p_{t,xy} V(y) = V(x) + \int_0^t e^{ks} \Big[ \sum_{z \in \mathbf{S}} p_{s,xz} \sum_{y \in \mathbf{S}} q_{zy} V(y) + k \sum_{y \in \mathbf{S}} p_{s,xy} V(y)\Big] ds.$$

Substituting Eq. (3.4) in the above expression, we have

$$\begin{aligned}
&e^{(k+\gamma)t}V(x)\sum_{y\in\mathbf{S}} p^V_{t,xy}\\
&= V(x)\Big(1+\int_0^t e^{(k+\gamma)s}\Big[\sum_{z\in\mathbf{S}} p^V_{s,xz}\sum_{y\in\mathbf{S}}(q^V_{zy}+\delta_{zy}\gamma)+k\sum_{y\in\mathbf{S}} p^V_{s,xy}\Big]ds\Big).
\end{aligned}$$

Dividing by $V(x)$, we obtain

$$e^{(k+\gamma)t}\sum_{y\in\mathbf{S}} p^V_{t,xy} = 1+\int_0^t e^{(k+\gamma)s}\Big[\sum_{z\in\mathbf{S}} p^V_{s,xz}\sum_{y\in\mathbf{S}} q^V_{zy}+\sum_{y\in\mathbf{S}}(k+\gamma)\,p^V_{s,xy}\Big]ds.$$

Now for $y=\Delta$, we directly have by the Kolmogorov forward equation:

$$e^{(k+\gamma)t}\,p^V_{t,x\Delta} = \int_0^t e^{(k+\gamma)s}\Big[p^V_{s,xz}\sum_{z\in\mathbf{S}} q^V_{z\Delta}+(k+\gamma)\,p^V_{s,x\Delta}\Big]ds.$$

Combining these, we obtain

$$\begin{aligned}
e^{(k+\gamma)t}\sum_{y\in\mathbf{S}_\Delta} p^V_{t,xy} &=1+\int_0^t e^{(k+\gamma)s}\Big[\sum_{z\in\mathbf{S}_\Delta} p^V_{s,xz}\sum_{y\in\mathbf{S}_\Delta} q^V_{zy}+\sum_{y\in\mathbf{S}_\Delta}(k+\gamma)\,p^V_{s,xy}\Big]ds\\
&=1+\int_0^t e^{(k+\gamma)s}\sum_{y\in\mathbf{S}_\Delta}(k+\gamma)\,p^V_{s,xy}ds\\
&\geq 1+\int_0^t e^{(k+\gamma)s}\sum_{y\in\mathbf{S}_\Delta}(k+\gamma)\,p^V_{t,xy}ds\\
&\geq 1+e^{(k+\gamma)t}\sum_{y\in\mathbf{S}_\Delta} p^V_{t,xy}-1\\
&=e^{(k+\gamma)t}\sum_{y\in\mathbf{S}_\Delta} p^V_{t,xy}.
\end{aligned}$$

The second equality is due to $Q^V$ being conservative. The inequality is due to non-increasingness of $s\mapsto\sum_{y\in\mathbf{S}_\Delta} p^V_{s,xy}$ (cf. [3, Proposition 1.1.2 (i)]). Since the first and last expressions are equal, the inequality is actually an equality. This yields that $\sum_{y\in\mathbf{S}_\Delta} p^V_{s,xy}$ is constant on $(0,t)$. Because $\sum_{y\in\mathbf{S}_\Delta} p^V_{s,xy}$ is continuous (cf. [3, Proposition 1.2.6]) it is also constant on $[0,t]$. Hence, $\sum_{y\in\mathbf{S}_\Delta} p^V_{s,xy}=1$ for $0\leq s\leq t$. From [3, Proposition 1.1.2(ii)] it follows that $\sum_{y\in\mathbf{S}_\Delta} p^V_{s,xy}=1$ for all $s\geq 0$. Hence $X^V$ is nonexplosive. □

By virtue of the above theorem for the $\gamma$-drift function $V$, requiring the nonexplosiveness of $X^V$ is necessary and sufficient for Eq. (3.3) to hold. Hence, Eq. (3.3) cannot hold for $V$ under weaker conditions.

In the next section we develop, in our opinion, satisfactory conditions implying the continuity properties on the parameter set $\mathsf{A}$ we will need.

## 3.3 Continuity for the parametrised processes

In order to address continuity aspects, we have to assume some structure on the parameter set.

**Assumption 3.3.1.** The set $\mathsf{A}$ is a locally compact topological space, in other words every point $a \in \mathsf{A}$ has a compact neighbourhood.

In what follows we will assume that the above condition holds.

**Definition 3.3.1.** $V : \mathbf{S} \to \mathbb{R}_+$ is called an $(\mathsf{A}, \gamma)$-drift function if $V$ is a $\gamma$-drift function for $X(a)$ for each $a \in \mathsf{A}$. The notions $(\mathsf{A}, \gamma)$-drift moment function and $(\mathsf{A}, \theta)$-drift $V$-moment function are defined accordingly. If the parameter space $\mathsf{A}$ consists of one element, we will drop the reference to $\mathsf{A}$ in the notation.

Recall the construction of the minimal transition function. Define

$$f_{t,xy}^{(n)}(a) = \begin{cases} \delta_{xy} e^{-q_x(a)t}, & n = 0 \\ f_{t,xy}^{(0)}(a) + \int_0^t e^{-q_x(a)} \sum_{k \neq x} q_{xk}(a) f_{t-s,ky}^{(n-1)}(a) ds, & n \geq 1. \end{cases}$$

By minimality of $X(a)$ [3, Theorem 2.2.2], one has

$$f_{t,xy}^{(n)}(a) \uparrow p_{t,xy}(a), \quad x, y \in \mathbf{S}, t \geq 0, a \in \mathsf{A}.$$

The interpretation is that $f_{t,xy}^{(n)}(a)$ is the probability that the process $X(a)$ reaches $y$ within $t$ time units with at most $n$ jumps when starting from state $x$.

**Theorem 3.3.1.** *Suppose that Assumptions 3.2.1 and 3.3.1 hold and that*

**i)** $a \mapsto q_{xy}(a)$ *is continuous on* $\mathsf{A}$ *for* $x, y \in \mathbf{S}$*;*

**ii)** *there exists an* $(\mathsf{A}, \gamma)$*-drift function* $V$*;*

**iii)** $(a, t) \mapsto P_t(a)V(x)$ *is continuous on* $\mathsf{A} \times [0, \infty)$*, for each* $x \in \mathbf{S}$*.*

*Then $(a,t) \mapsto p^V_{t,xy}(a)$ continuous on $\mathsf{A} \times [0,\infty)$ for each $x,y \in \mathbf{S}$. Hence, $(a,t) \mapsto p_{t,xy}(a)$ is continuous on $\mathsf{A} \times [0,\infty)$ for each $x,y \in \mathbf{S}$.*

*Proof.* Let $f^{V,(n)}_{t,xy}(a)$ be the above probabilities for the $V$-transformed process $X^V(a)$. Thus,

$$f^{V,(n)}_{t,xy}(a) = \begin{cases} \delta_{xy} e^{-q^V_x(a)t}, & n = 0 \\ f^{V,(0)}_{t,xy}(a) + \int_0^t e^{-q^V_x(a)} \sum_{k\neq x} q^V_{xk}(a) f^{V,(n-1)}_{t-s,ky}(a) ds, & n \geq 1. \end{cases}$$

We will inductively show that $(a,t) \mapsto \sum_{y\in K} f^{V,(n)}_{t,xy}(a)$ is continuous for each $n \geq 1$, $x \in \mathbf{S}_\Delta$, $K \subset \mathbf{S}_\Delta$. First we will show this statement for $K = \{y\}$. Note that $(a,t) \mapsto f^{V,(0)}_{t,xy}(a) = e^{-q^V_x(a)t}\delta_{xy}$ is continuous for $x,y \in \mathbf{S}_\Delta$.

Assume that $(a,t) \mapsto f^{V,(n-1)}_{t,xy}(a)$ is continuous for each, $x,y \in \mathbf{S}_\Delta$. Because $f^{V,(n-1)}_{t,xy}(a) \leq 1$, the generalised dominated convergence theorem [58, Proposition 11.18] implies that $(a,t) \mapsto \sum_{k\neq x} q^V_{xk}(a) f^{V,(n-1)}_{t,ky}(a)$ is continuous for each $x,y \in \mathbf{S}_\Delta$. For each $(a,t) \in \mathsf{A} \times [0,\infty)$ this expression is bounded by $q^V_x(a)$. Applying the generalised dominated convergence theorem once more yields that the integral $\int_0^t e^{-q^V_x(a)} \sum_{k\neq x} q^V_{xk}(a) f^{V,(n-1)}_{t-s,ky}(a) ds$ is a continuous function of $(a,t) \in \mathsf{A} \times [0,\infty)$. This gives continuity of $(a,t) \mapsto f^{V,(n)}_{t,xy}(a)$, for $x,y \in \mathbf{S}_\Delta$.

An analogous argument shows continuity of $(a,t) \mapsto \sum_{y\in K} f^{V,(n)}_{t,xy}(a)$ for any subset $K \subset \mathbf{S}_\Delta$, $x \in \mathbf{S}_\Delta$. By virtue of Eq. (3.4), Theorem 3.3.1(iii) is equivalent to requiring continuity of $(a,t) \mapsto \sum_{y\in\mathbf{S}} p^V_{t,xy}(a)$.

Next, let $x,y \in \mathbf{S}$. We wish to show that $(a,t) \mapsto p^V_{t,xy}(a)$ is continuous at some arbitrary point $(a_0,t_0) \in \mathsf{A} \times [0,\infty)$. Let $B_0 \subset \mathsf{A} \times [0,\infty)$ be a compact neighbourhood of $(a_0,t_0)$. Hence, $(a,t) \mapsto \sum_{y\in\mathbf{S}} f^{V,(n)}_{t,xy}(a)$ is a nondecreasing sequence of continuous functions on a compact set, converging to the (assumed) continuous function $(a,t) \mapsto \sum_{y\in\mathbf{S}} p^V_{t,xy}(a)$. By Dini's theorem on uniform convergence [59, Theorem 7.13], the convergence is uniform. In other words, for each $\epsilon > 0$, there exists $N_\epsilon$ such that for $(a,t) \in B_0, n \geq N_\epsilon$

$$\epsilon \geq |\sum_{y\in\mathbf{S}} p^V_{t,xy}(a) - \sum_{y\in\mathbf{S}} f^{V,(n)}_{t,xy}(a)| = \sum_{y\in\mathbf{S}} \big( p^V_{t,xy}(a) - f^{V,(n)}_{t,xy}(a) \big).$$

As a consequence, $f^{V,(n)}_{t,xy}(a)$ converges uniformly in $(a,t) \in B_0$ to $p^V_{t,xy}(a)$, for $x,y \in \mathbf{S}$, $t \geq 0$.

By virtue of the uniform limit theorem (cf. [50, Theorem 21.6, p. 132] and [58, Exercise 2.42]), $(a,t) \mapsto p^V_{t,xy}(a)$ is continuous in $(a_0,t_0)$, for $x,y \in \mathbf{S}$. Continuity of $(a,t) \mapsto p_{t,xy}(a)$ then follows by Eq. (3.4). $\square$

**Corollary 3.3.2.** *Suppose that Assumptions 3.2.1 and 3.3.1 hold and that $X(a)$ is nonexplosive for all $a \in \mathsf{A}$. Furthermore, assume that $a \mapsto q_{xy}(a)$ is continuous for each $x, y \in \mathbf{S}$. Then $(a,t) \mapsto p_{t,xy}(a)$ is continuous for $x, y \in \mathbf{S}$.*

*Proof.* It holds that $V \equiv 1$ is always a 0-drift function. Furthermore, we have $P_t(a)V(x) = 1$, $x \in \mathbf{S}$; hence, $(a,t) \mapsto P_t(a)V$ is continuous on $\mathbf{S} \times [0,\infty)$. The result follows from the previous theorem. □

Clearly, Theorem 3.3.1(iii) is not easily verified for general drift functions. The next theorem provides verifiable conditions for the conditions of Theorem 3.3.1 to hold.

Simultaneously with the preparation of this work, this question has been addressed in [53, Proposition 2.20]. Due to the equivalence result of Theorem 3.2.1, the result in [53] is close to ours. The book [53], however, restricts to a product set parameter space, and requires compactness of the parameter space. We will provide an alternative proof. The conditions required are the existence of a $\gamma$-drift function $V$ and $\theta$-drift $V$-moment function $W$, uniform in the parameter $a \in \mathsf{A}$.

**Assumption 3.3.2.** **i)** It holds that $a \mapsto q_{xy}(a)$ is continuous on $\mathsf{A}$ for $x, y \in \mathbf{S}$.

**ii)** There exists a $(\mathsf{A},\gamma)$-drift function $V$.

**iii)** There exists a $(\mathsf{A},\theta)$-drift $V$-moment function $W$.

**Theorem 3.3.3.** *Suppose that Assumptions 3.2.1, 3.3.1 and 3.3.2 hold. Then for each $x \in \mathbf{S}$, $(a,t) \mapsto P_t(a)V(x)$ is continuous on $\mathsf{A} \times [0,\infty)$ and $a \mapsto Q(a)V(x)$ is continuous on $\mathsf{A}$.*

*Proof.* Denote by $P_t^V(a)$, $t \geq 0$, the transition function of $X^V(a)$. Since $X^V(a)$ is non-explosive by virtue of Theorem 3.2.1,

$$\sum_{y \in \mathbf{S}_\Delta} p^V_{t,xy}(a) = 1, \qquad \text{for all } (a,t) \in \mathsf{A} \times [0,\infty), x \in \mathbf{S}_\Delta.$$

Moreover Corollary 3.3.2 yields that $(a,t) \mapsto p^V_{t,xy}(a)$ is continuous for $x, y \in \mathbf{S}_\Delta$. Combining this gives continuity of

$$(a,t) \mapsto \sum_{y \in \mathbf{S}} p^V_{t,xy}(a) \tag{3.5}$$

on $\mathsf{A} \times [0, \infty)$, for $x \in \mathbf{S}$. Inserting Eq. (3.4) in Eq. (3.5) yields continuity of $a \mapsto \sum_{y \in \mathbf{S}} p_{t,xy}(a) V(y)$ for each $x \in \mathbf{S}$.

The only thing left to prove is continuity of $a \mapsto Q(a)V(x)$. To this end we use a nice argument used to prove [53, Proposition 2.20]. Let $x \in \mathbf{S}$ be given. Let $\{K_n\}_n \subset \mathbf{S}$ be an increasing sequence of finite sets with $x \in K_n$ for all $n$, $\lim_n K_n = \mathbf{S}$, and $\inf_{y \notin K_n} W(y)/V(y) \to \infty$ as $n \to \infty$. Then, for all $a \in \mathsf{A}$,

$$\begin{aligned}
\sum_{y \notin K_n} q_{xy}(a) V(y) &= \sum_{y \notin K_n} q_{xy}(a) W(y) \frac{V(y)}{W(y)} \\
&\leq \frac{1}{\inf_{z \notin K_n} W(z)/V(z)} \sum_{y \notin K_n} q_{xy}(a) W(y) \\
&\leq \frac{1}{\inf_{z \notin K_n} W(z)/V(z)} (\theta + q_x(a)) W(x).
\end{aligned}$$

Let $a_0 \in \mathsf{A}$. We wish to show that $a \mapsto \sum_y q_{xy}(a) V(y)$ is continuous in $a_0$. Let $\mathsf{A}_0$ be a compact neighbourhood of $a_0$. Then $b := \sup_{a \in \mathsf{A}_0} (\theta + q_x(a)) W(x) < \infty$. For any $\epsilon > 0$ there exists $N_\epsilon$, such that

$$\frac{b}{\inf_{z \notin K_n} W(z)/V(z)} \leq \epsilon, \quad n \geq N_\epsilon.$$

It follows that $\sum_{y \in K_n} q_{xy}(a) V(y)$ converges to $Q(a)V(x)$, uniformly in $a \in \mathsf{A}_0$. Since $a \mapsto \sum_{y \in K_n} q_{xy}(a) V(y)$ is continuous by assumption, we may apply the uniform limit theorem (cf. [50, Theorem 21.6, p. 132]) to obtain that $a \mapsto Q(a)V(x)$ is continuous. $\square$

Theorem 3.3.3 links continuity properties of the integrals $(a,t) \mapsto P_t(a) f(x)$ for $f \in \ell^\infty(\mathbf{S}, V)$, to continuity of the measures of compact sets and non-explosiveness properties of $X(a)$. The next example illustrates that if Assumption 3.3.2 (i,ii) hold but $X^V(a)$ is explosive for some $a \in \mathsf{A}$, then $a \mapsto P_t(a)V(x)$ need not be continuous on $\mathsf{A}$. This is the basic example from [66, Section 4].

**Example 3.3.1.** Let $\mathbf{S} = \mathbf{Z}_+$. Consider the $q$-matrix $Q$ given by

$$q_{xy} = \begin{cases} p2^x, & y = x+1, x \neq 0 \\ -2^x, & y = x, x \neq 0 \\ (1-p)2^x, & y = x-1, x \neq 0 \\ 0, & \text{else,} \end{cases}$$

where $p < 1/2$. 0 is an absorbing state. This is the $q$-matrix of a nonexplosive Markov process.

Let $V(x) = \alpha^x$, for $\alpha = (1-p)/p$. Then $QV = 0 \leq 0 \cdot V$. The $q$-matrix $Q^V$ of the associated $V$-transformation however defines an explosive Markov process $X$ (cf. [66]).

We define the following parametrised collection of Markov processes. Let $\mathsf{A} = \{1, 2, \ldots, \infty\}$. This is a compact set. Let $X(0)$ be the Markov process with $q$-matrix $Q(\infty) = Q$. For each $a \in \mathsf{A}$ we define the perturbation $X(a)$ with $q$-matrix $Q(a)$ given by

$$q_{xy}(a) = \begin{cases} q_{xy}, & x \leq a \\ -2^a, & y = x > a \\ 2^a, & y = x-1, x > a. \end{cases}$$

Then $Q(a)V \leq 0 \cdot V$ for every $a \in \mathsf{A}$. Also $a \mapsto q_{xy}(a)$ is trivially continuous on $\mathsf{A}$. Hence, Assumptions 3.3.2 (i) and 3.3.2 (ii) are satisfied. Note that due to the boundedness of jumps, $X^V(a)$, $a < \infty$, is nonexplosive.

Since $X^V = X^V(\infty)$ is explosive, there exists a state $x \in \mathbf{S}_\Delta$ such that

$$1 = \lim_{a\to\infty} \sum_y p^V_{t,x,y}(a) > \sum_y p^V_{t,x,y}(\infty).$$

By virtue of Eq. (3.4), $\sum_y p_{t,xy}(a)V(y) \not\to \sum_y p_{t,xy}(\infty)V(y)$ as $a \to \infty$, for $t > 0$. Hence, $a \mapsto P_t(a)V(x)$ is not continuous on $\mathsf{A}$.

## 3.4 Poisson and optimality equation for the $\alpha$-discounted cost criterion

Suppose next that Assumptions 3.2.1 and 3.3.2 (ii) hold, in other words there exists a $\gamma$-drift function $V$. Assume that a cost $c_x(a)$ per unit time is incurred when the process $X(a)$ resides in state $x$ under parameter $a \in \mathsf{A}$. Denote by $c(a) = (c_x(a))_{x\in\mathbf{S}}$ the associated cost vector.

**Assumption 3.4.1. i)** It holds that $a \mapsto c_x(a)$ is continuous on $\mathsf{A}$.

**ii)** There is a finite constant $c_V$ such that $\sup_{x,a} |c_x(a)|/V(x) \leq c_V$.

**iii)** For the discount factor $\alpha$ it holds that $\alpha > \gamma$.

Define the expected $\alpha$-discount total cost associated with parameter $a \in \mathsf{A}$ by

$$v^\alpha(a) = \int_0^\infty e^{-\alpha t}\, P_t(a)c(a)dt,$$

and the $x$th component by $v^\alpha(x,a)$. Note that Assumptions 3.2.1, 3.3.2(ii) and 3.4.1(ii) imply that $t \mapsto P_t(a)V$ is continuous. By Eq. (3.4) $P_tV(x) \le e^{\gamma t}V(x)$. Hence $\alpha > \gamma$ guarantees that $v^\alpha(a)$ is well-defined and finite.

If we further require nonexplosiveness of $X^V$, then it can be shown that $v^\alpha(a)$ is the unique solution of the Poisson equation Eq. (3.6) in $\ell^\infty(\mathbf{S}, V)$.

**Theorem 3.4.1.** *Let Assumption 3.2.1 hold.*

*i) If Assumptions 3.3.2(ii) and 3.3.2(iii), and Assumptions 3.4.1(ii) and 3.4.1 (iii) hold, then $v^\alpha(a)$ is the unique solution in $\ell^\infty(\mathbf{S}, V)$ to the $\alpha$-discounted equation*

$$\alpha f = c(a) + Q(a)f. \tag{3.6}$$

*ii) If, additionally, Assumptions 3.3.1, 3.3.2(i) and 3.4.1(i) hold, then $a \mapsto v^\alpha(a)$ is component-wise continuous on $\mathsf{A}$.*

*Proof.* Let $a \in \mathsf{A}$. We first prove that $v^\alpha(a)$ is a solution to Eq. (3.6) in the space $\ell^\infty(\mathbf{S}, V)$. Note that $\|v^\alpha(a)\|_V \le c_v/(\alpha-\gamma)$, so that $v^\alpha(a) \in \ell^\infty(\mathbf{S}, V)$. Moreover $Q(a)|v^\alpha(a)|$ is well defined and finite. We obtain

$$\begin{aligned}
Q(a)v^\alpha(x,a) &= \sum_y q_{xy}(a)\int_0^\infty e^{-\alpha t}\sum_z p_{t,yz}(a)c_z(a)dt \\
&= \sum_y q_{xy}(a)\sum_z \int_0^\infty e^{-\alpha t}\, p_{t,yz}(a)dt\, c_z(a) \\
&= \sum_z \int_0^\infty e^{-\alpha t}\, p'_{t,xz}(a)dt\, c_z(a) \\
&= \sum_z \Big(-\delta_{xz} + \alpha\int_0^\infty e^{-\alpha t}\, p_{t,xz}(a)dt\Big)\, c_z(a) \\
&= -c_x(a) + \alpha v^\alpha(x,a).
\end{aligned}$$

The interchange of summation and integration in the second equality is justified by Fubini's theorem; in the third equality by the additional fact that $Q(a)$ has at most one negative element per row. The fourth equality is due to partial integration. As a consequence, $v^\alpha(a)$ is a solution of Eq. (3.6) in $\ell^\infty(\mathbf{S}, V)$.

Suppose that $f \in \ell^\infty(\mathbf{S}, V)$ is another solution. Then $\alpha(v^\alpha(a) - f) = Q(a)(v^\alpha(a) - f)$, and so $v^\alpha(a) - f \in \ell^\infty(\mathbf{S}, V)$ is an eigenvector of $Q(a)$ to eigenvalue $\alpha > 0$. Direct calculations show that $g : \mathbf{S}_\Delta \to \mathbb{R}$ given by

$g = (v^\alpha(a) - f)/V$ on $\mathbf{S}$ and $g(\Delta) = 0$ is a bounded eigenvector of $Q^V(a)$ to eigenvalue $\alpha - \gamma > 0$. Nonexplosiveness of a Markov process can be characterised by the nonexistence of a bounded (nonzero) eigenvector of the corresponding $q$-matrix to positive eigenvalues (cf. [55, Theorem 7] and [67, Theorem 2.1]). By Assumption 3.3.2, $X^V(a)$ is nonexplosive. Hence a nonzero eigenvector cannot exist and so we conclude that $f = v^\alpha(a)$.

We finally turn to proving component-wise continuity of $v^\alpha(a)$. By virtue of Theorem 3.3.3, $a \mapsto P_t(a)V$ is component-wise continuous. Eq. (3.2) yields that $P_t(a)V \le e^{\gamma t}V$. Hence, the dominated convergence theorem implies that $a \mapsto \int_0^\infty e^{-\alpha t} P_t(a)V dt$ is componentwise continuous. Another application of the dominated convergence theorem implies that $a \mapsto v^\alpha(a) = \int_0^\infty e^{-\alpha t} P_t(a)c(a)dt$ is componentwise continuous. $\square$

We will next consider the special case that the sets $\{Q(a)\}_{a\in\mathsf{A}}$ and $\{c(a)\}_{a\in\mathsf{A}}$ have the product property (cf. [39]) in the following sense.

**Assumption 3.4.2.** There exist compact metric sets $\mathsf{A}_x$, $x \in \mathbf{S}$, such that the following conditions hold:

**i)** $\mathsf{A} = \prod_{x\in\mathbf{S}} \mathsf{A}_x$, and $\mathsf{A}$ is equipped with the product topology;

**ii)** $\{Q(a)\}_{a\in\mathsf{A}}$ and $\{c(a)\}_{a\in\mathsf{A}}$ have the product property. In other words, for any $a, a' \in \mathsf{A}$, $x \in \mathbf{S}$ such that that $a_x = a'_x$, it holds that $(Q(a))_{x\cdot} = (Q(a'))_{x\cdot}$, and $c_x(a) = c_x(a')$. Here $(Q(a))_{x\cdot}$ stands for the $x$-row of $Q(a)$.

Note that $\mathsf{A}$ is compact and metrisable, and the product topology is the topology of component-wise convergence. Hence $\mathsf{A}$ is sequentially compact.

Under Assumption 3.4.2 the $x$th row and $x$th component of $Q(a)$ and $c(a)$ depend on the value $a_x$ only. Therefore, with a slight abuse of notation, we may write $q_{xy}(a_x)$ and $c_x(a_x)$. Then $\inf_{a\in\mathsf{A}}\{c(a) + Q(a)f\}$ is well defined and may also be written as $\inf_{a_x\in\mathsf{A}_x}\{c_x(a_x) + \sum_y q_{xy}(a_x)f(y)\}$, for all $x \in \mathbf{S}$. As an application, the set $\mathsf{A}$ may represent the collection of stationary policies in an MDP, or the set of deterministic policies.

We say that *parameter $a^*$ is optimal in $\mathsf{A}$ if $v^\alpha(a^*) \le v^\alpha(a)$ for all $a \in \mathsf{A}$.* In this case we have the following result.

**Theorem 3.4.2.** *Suppose that Assumption 3.2.1 holds.*

*i) Suppose that Assumptions 3.3.2(ii), 3.4.1(ii), 3.4.1(iii), and 3.4.2 hold. Moreover suppose there exists a function $m$ such that $m(x) \ge \sup_a q_x(a)$.*

*Then the equation*

$$\alpha f(x) = \inf_{a_x \in \mathsf{A}_x} \{c_x(a_x) + \sum_y q_{xy}(a_x) f(y)\}, \quad x \in \mathbf{S}, \tag{3.7}$$

*has a solution* $v^\alpha$ *in* $\ell^\infty(\mathbf{S}, V)$.

*ii) If, moreover, Assumptions 3.3.2(i), 3.3.2(iii) and 3.4.1(i) hold, this solution is unique in* $\ell^\infty(\mathbf{S}, V)$ *and the infimum is a minimum. For any* $a^* = (a^*_x)_x \in \mathsf{A}$ *for which* $a^*_x$ *achieves the minimum in Eq. (3.7) ,* $x \in \mathbf{S}$, *one has* $v^\alpha(a^*) = v^\alpha$ *and* $a^*$ *is optimal in* $\mathsf{A}$.

*Proof of Theorem 3.4.2(i).* We use the same line of reasoning as in the proof of [53, Theorem 3.7]. Suppose that Assumptions 3.3.2(ii), 3.4.1(ii), and 3.4.2 hold. Let $m : \mathbf{S} \to \mathbb{R}_+$ be such that $m(x) \geq \sup_{a_x \in A_x} q_x(a_x)$, for $x \in \mathbf{S}$. Then define $p_{xy}(a_x) = q_{xy}(a_x)/m(x) + \delta_{xy}$ for $x, y \in \mathbf{S}, a_x \in A_x$, which is a probability measure for each state action pair $(x, a_x)$. Furthermore, define the operator $T$ for $f \in \ell^\infty(\mathbf{S}, V)$ by

$$(Tf)(x) = \inf_{a_x \in \mathsf{A}_x} \left\{ \frac{c_x(a_x)}{\alpha + m(x)} + \frac{m(x)}{\alpha + m(x)} \sum_{y \in \mathbf{S}} p_{xy}(a_x) f(y) \right\}, \quad x \in \mathbf{S}.$$

Define the sequence $\{f_n\}_n$ in $\ell^\infty(\mathbf{S}, V)$ by $f_0(x) = (c_V(\alpha - \gamma))V(x)$, and $f_n = Tf_{n-1}$ for $n \geq 1$. First, nonnegativity of the coefficients in the second term between brackets implies that $T$ is monotone (i.e. $f \geq g$ implies that $Tf \geq Tg$). Secondly, direct calculations show that $f_0 \geq f_1$. This implies that $\{f_n\}_n$ is a monotone decreasing sequence. Further it is easy to show that

$$\|f_n\|_V \leq \frac{c_V}{\alpha - \gamma}.$$

Thus, $\{f_n\}_n$ has a pointwise limit $f^* \in \ell^\infty(\mathbf{S}, V)$ with $f^* \leq f_n$ for all $n$. Hence, $Tf^* \leq Tf_n = f_{n+1}$ for all $n$, and, thus, $Tf^* \leq \lim_{n\to\infty} f_n = f^*$.

Next we prove that $f^* \leq Tf^*$. First note that

$$f^* \leq f_{n+1} = Tf_n, \quad n = 1, \ldots. \tag{3.8}$$

For notational convenience, denote

$$(T_{a_x} f)(x) = \frac{c_x(a_x)}{\alpha + m(x)} + \frac{m(x)}{\alpha + m(x)} \sum_{y \in \mathbf{S}} p_{xy}(a_x) f(y), \quad x \in \mathbf{S},$$

so $Tf(x) = \inf_{a_x} T_{a_x} f(x)$. By monotone convergence $T_{a_x} f_n(x) \downarrow T_{a_x} f^*(x)$, $n \to \infty$, $a_x \in \mathsf{A}_x$. Let $\epsilon > 0$, $x \in \mathbf{S}$, and $a_x \in \mathsf{A}_x$. Then there exists $N_{\epsilon,x,a_x}$ such that

$$T_{a_x} f_n(x) \le T_{a_x} f^*(x) + \epsilon, \quad n \ge N_{\epsilon,x,a_x}. \tag{3.9}$$

Combining Eq. (3.9) with Eq. (3.8) yields

$$f^*(x) \le T_{a_x} f^*(x) + \epsilon, \quad a_x \in \mathsf{A}_x.$$

Taking the infimum on both sides gives

$$f^*(x) \le T f^*(x) + \epsilon.$$

Since $\epsilon > 0$ and $x \in \mathbf{S}$ were arbitrary, we get the desired inequality $f^* \le T f^*$. We conclude that $T f^* = f^*$.

By direct calculations it is seen that this last equality is equivalent to Eqn. (3.7); thus, we have proven that there is a solution and we call this $v^\alpha$.

*Proof of Theorem 3.4.2(ii).* Suppose now that Assumptions 3.3.2(i), 3.3.2(iii), and 3.4.1(i) hold as well. By Theorem 3.6, $a \mapsto c(a) + Q(a)v^\alpha$ is componentwise continuous on $\mathsf{A}$. Since $\mathsf{A}$ is compact, this implies that the infimum is attained. So there is an $a^* \in \mathsf{A}$ such that

$$\begin{aligned} \alpha v^\alpha(x) &= \inf_{a_x \in \mathsf{A}_x} \{c_x(a_x) + \sum_y q_{xy}(a_x) v^\alpha(y)\} \\ &= \min_{a_x \in \mathsf{A}_x} \{c_x(a_x) + \sum_y q_{xy}(a_x) v^\alpha(y)\} \\ &= c_x(a_x^*) + \sum_y q_{xy}(a_x^*) v^\alpha(y). \end{aligned}$$

Then $v^\alpha = v^\alpha(a^*)$ by Theorem 3.4.1. Next we will show that $v^\alpha = v^\alpha(a^*) \le v^\alpha(a)$ for any $a \in \mathsf{A}$ in other words, $a^*$ is optimal in $\mathsf{A}$. To this end, let $\hat{a} \in \mathsf{A}$. Enumerate $\mathbf{S} = \{s_1, s_2, \ldots\}$. Define $a^n \in \mathsf{A}$ by

$$a_x^n = \begin{cases} \hat{a}_x, & x \in \{s_1, \ldots, s_n\} \\ a_x^*, & x \in \{s_{n+1}, \ldots\}. \end{cases}$$

Then $a^n \to \hat{a}$, $n \to \infty$, in the product topology and, in particular,

$$\alpha v^\alpha \le c(a^n) + Q(a^n) v^\alpha.$$

Define

$$d^n = c(a^n) + Q(a^n)v^\alpha - \alpha v^\alpha,$$

then $d^n$ has at most $n$ nonzero components and so $d^n \in \ell^\infty(\mathbf{S}, V)$. It follows that $|Q(a^n)v^\alpha| \in \ell^\infty(\mathbf{S}, V)$. Hence, $t \mapsto P_t(a^n)(Q(a^n)v^\alpha)$ is finite and continuous, and

$$\alpha\, P_t(a^n)v^\alpha \leq P_t(a^n)c(a^n) + P_t(a^n)(Q(a^n)v^\alpha).$$

Multiplying both sides by $e^{-\alpha t}$, integrating over $(0, T)$, and rearranging terms, we obtain, for any $T > 0$,

$$\int_0^T e^{-\alpha t}\big[\alpha\, P_t(a^n)v^\alpha - P_t(a^n)(Q(a^n)v^\alpha)\big]dt \leq \int_0^T e^{-\alpha t}\, P_t(a^n)c(a^n)dt.$$

By virtue of Theorem 3.2.2, Eq. (3.3) is applicable with $k = -\alpha$, thus yielding

$$v^\alpha - e^{-\alpha T}\, P_T(a^n)v^\alpha \leq \int_0^T e^{-\alpha t}\, P_t(a^n)c(a^n)dt, \quad T > 0.$$

Note that $\|\, P_T(a^n)v^\alpha\|_V \leq e^{\gamma T}\|v^\alpha\|_V \leq e^{\gamma T}c_V/(\alpha - \gamma)$. Taking the limit $T \to \infty$, we obtain the desired result that $v^\alpha \leq v^\alpha(a^n)$. Since $a \mapsto v^\alpha(a)$ is component-wise continuous, we can finally take the limit $n \to \infty$ and obtain that $v^\alpha \leq v^\alpha(\hat{a})$. Uniqueness now follows immediately. □

*Remark* 3.4.1. The question arises whether an optimal policy in $\mathsf{A}$ is optimal in the class of *Markov policies*, as defined in [53], or even in more general classes of policies. Notice that a Markov policy generates a nonhomogeneous Markov process. Following the proof that a solution to the $\alpha$-discount optimality equation dominates the expected $\alpha$-discounted cost under a Markov policy in [53, Lemma 3.5], one needs the result of Theorem 3.2.2 to hold for a nonhomogeneous Markov process. To our knowledge, such a result has not yet been formally proved.

*Discussion on related conditions in the literature.* In [34], [36] and [53] the parametrised process $X(a)$ as well as $X^V(a)$ are supposed to be nonexplosive for all $a \in \mathsf{A}$. We only require $X^V(a)$ to be nonexplosive uniformly in $\mathsf{A}$. This relaxation might be useful if the cost function goes to zero 'fast enough' as the state grows large. See the example below, it is a variation on Example 3.3.1. In [51] $X(a)$ is not required to be nonexplosive either, however, the extra condition that $q_x(a)V(x) \leq W(x)$ for $x \in \mathbf{S}$, $a \in \mathsf{A}$, is required there. In [67] a detailed discussion on the relation between the various drift conditions used in this context is presented.

**Example 3.4.1.** Let $\mathbf{S} = \mathbf{Z}_+$. Define the following $q$-matrices $Q(a)$ by

$$q_{xy}(a) = \begin{cases} a_x 2^x, & y = x+1, x \neq 0 \\ -2^x, & y = x, x \neq 0 \\ (1-a_x)2^x, & y = x-1, x \neq 0 \\ 0, & \text{else}, \end{cases}$$

for any $a_x \in \mathsf{A}_x = [p_0, p_1]$, with $\frac{1}{2} < p_0 \leq p_1 \leq 1$. Hence $\mathsf{A} = \prod_{x \in \mathbf{S}} \mathsf{A}_x$ is a compact product set. Notice that clearly $a \mapsto q_{xy}(a)$ is continuous on $\mathsf{A}$. Notice also that since $a_x \geq p_0 > 1/2$ for all $a_x$, this is an explosive Markov process, for every $a \in \mathsf{A}$.

Next define the reward structure $r(a)$ (note that nowhere in the theory above is it essential whether to maximise or minimise). We let the reward rate consist of two parts: a fixed reward rate $B$ for staying in the finite set $\{x \leq U\}$, and a bonus depending on the current state for taking actions that move the system to a higher state with larger probability. Therefore, put $r_x(a) = b_x(a) + c_x(a)$, with $b_x(a) = B\mathbb{1}_{\{x \leq U\}}$ and $c_x(a) = Ca_x(1-\epsilon)^x$, where $\mathbb{1}$ is the indicator function.

We will make a transformation that makes the transformed process nonexplosive, take $V(x) = \beta^x$, with $\max\left((1-p_0)/p_0, 1-\epsilon\right) \leq \beta < 1$. Notice that since $\beta < 1$, $V$ is not a moment function. Then for all $a \in \mathsf{A}, x \in \mathbf{S}$

$$\begin{aligned} Q(a)V(x) &= \left(a_x\beta^{x+1} - \beta^x + (1-a_x)\beta^{x-1}\right)2^x \\ &= \left(a_x\beta^2 - \beta + (1-a_x)\right)\beta^{x-1}2^x \\ &= \left(a_x(\beta - \frac{1-a_x}{a_x})(\beta-1)\right)\beta^{x-1}2^x \\ &\leq 0 \cdot V, \end{aligned}$$

the inequality holds because $(1-a_x)/a_x < \beta \leq 1$. Hence, $V$ is a $(\mathsf{A}, 0)$-drift function. Moreover, $r(a)$ is uniformly $V$-bounded, since

$$\sup_{x \in \mathbf{Z}_+, a \in \mathsf{A}} \frac{|b_x(a)|}{V(x)} = \max_{x \leq U} B\beta^{-x} = B\beta^{-U},$$

and

$$\sup_{x \in \mathbf{Z}_+, a \in \mathsf{A}} \frac{|c_x(a)|}{V(x)} = \max_{x \in \mathbf{Z}_+} Cp_1 \left(\frac{1-\epsilon}{\beta}\right)^x = Cp_1.$$

Next take $W \equiv 1$, then $Q(a)W = 0 \leq 0 \cdot W$ and $\lim_{x\to\infty} W(x)/V(x) = \lim_{x\to\infty} \beta^{-x} = \infty$. Hence $W$ is a $(\mathsf{A}, 0)$-drift $V$-moment function. Then Theorem 3.3.3 yields the transformed process $X^V(a)$ is non-explosive for all $a \in \mathsf{A}$.

Now all assumptions of Theorem 3.4.2 hold, hence, (3.7) (with the infimum replaced by a supremum) has a unique solution $v^\alpha \in \ell^\infty(\mathbf{S}, V)$ for any $\alpha > 0$ and there is a parameter $a^* \in \mathsf{A}$ that achieves this supremum.

## 3.5 MDPs and perturbations

In this section we show how Theorem 3.4.2 can be applied to MDPs. In order to do so, we take the parameter set $\mathsf{A} := \mathcal{D} = \prod_x \mathcal{D}_x$, where $\mathcal{D}$ is the set of all deterministic (stationary) policies and $\mathcal{D}_x = \{\text{set of actions available in state } x\}$, $x \in \mathbf{S}$. Then $\mathsf{A} = \mathcal{D}$ has the product property described in Assumption 3.4.2. We use the notation $\delta \in \mathcal{D}$ for a deterministic (stationary) policy and by $\delta(x) \in \mathcal{D}_x$ the corresponding action prescribed in state $x$ by $\delta$. If we assume that $\mathcal{D}_x$ is a compact, metric space for each $x \in \mathbf{S}$, then $\mathcal{D}$ is a compact, metric space as well. Consequently, an MDP with compact action space and deterministic policies $\mathcal{D}$, can be identified with a parametrised collection of Markov processes satisfying Assumption 3.4.2.

*Remark* 3.5.1. If Assumptions 3.2.1, 3.3.2, 3.4.1, and 3.4.2 hold for $\mathsf{A} = \mathcal{D}$, it is a standard construction to show that these assumptions apply as well for the parameter set equal to the set $\mathcal{S}$ of stationary, randomised policies. For an example of this construction see [29]. Hence, the assertion of Theorem 3.4.2 then also applies for this larger parameter set. Furthermore, it is a simple consequence that if $\mathcal{A} = \mathcal{S}$ in Eq. (3.7) there exists a minimiser $\delta^* \in \mathcal{D}$ for which $v^\alpha(\delta^*) = v^\alpha$. As a consequence, we may (and we will) restrict our analysis to $\mathcal{D}$.

**Perturbation of MDPs.** In this paragraph we will discuss how Theorems 3.4.1 and 3.4.2 can be applied to analyse MDPs by adding a perturbation.

The application we have in mind is the analysis of structural properties of an MDP with unbounded transition rates (i.e. $\sup_{x \in \mathbf{S}, \delta \in \mathcal{D}} q_x(\delta) = \infty$), and, thus, the uniformisation technique is not applicable. In particular, we are interested in the structure of optimal strategies and of the value function. To this end we perturb the MDP to get bounded rates so that it can be studied using the discrete time equivalent MDP. This perturbation is indexed by an extra parameter $N$, typically $N \in \mathcal{N}$, where $\mathcal{N} := \{1, 2, \ldots, \infty\}$, a compact set. Thus we obtain a collection of extended parametrised processes, $\{X(N, \delta)\}_{(N,\delta) \in \mathcal{N} \times \mathcal{D}}$. For fixed $N$ the parametrised process $\{X(N, \delta)\}_{\delta \in \mathcal{D}}$ is an MDP and for $N = \infty$ this coincides with the original MDP. The theorems in the previous section provide the framework that guarantees continuity in the perturbation parameter. This induces convergence of the results for the perturbed models to the original model if the perturbation vanishes, i.e. the parameter goes to infinity.

**Theorem 3.5.1.** *Consider an MDP, in other words a parametrised collection of processes $\{X(\delta)\}_{\delta\in\mathcal{D}}$ with cost function $c(\delta)_{\delta\in\mathcal{D}}$. Furthermore, consider an extended parametrised collection of processes $\{X(N,\delta)\}_{(N,\delta)\in\mathcal{N}\times\mathcal{D}}$ with cost function $c(N,\delta)_{(N,\delta)\in\mathcal{N}\times\mathcal{D}}$ such that $X(\infty,\delta)=X(\delta)$ and $c(\infty,\delta)=c(\delta)$.*

*Suppose that Assumptions 3.2.1, 3.3.1, 3.3.2, and 3.4.1 hold for the collection $\{X(N,\delta)\}_{(N,\delta)\in\mathcal{N}\times\mathcal{D}}$. Suppose that additionally Assumption 3.4.2 holds for $\{X(N,\delta)\}_{\delta\in\mathcal{D}}$ for all $N\in\mathcal{N}$. Let $v^\alpha_N$ be the value function for the MDP $\{X(N,\delta)\}_{\delta\in\mathcal{D}}$ and $\delta^*_N$ an optimal policy, $N\in\mathcal{N}$. Then the following hold:*

*i)* $\lim_{N\to\infty} v^\alpha_N = v^\alpha$*;*

*ii) any limit point of* $(\delta^*_N)_{N\in\mathcal{N}}$ *is optimal for* $\{X(\delta)\}_{\delta\in\mathcal{D}}$.

*Proof.* The assertions of Theorem 3.4.2 hold for $\{X(N,\delta)\}_{\delta\in\mathcal{D}}$ for fixed $N\in\mathcal{N}$. This yields the existence of a pair $(v^\alpha_N,\delta^*_N)$ satisfying Eq. (3.7), so that $v^\alpha_N=v^\alpha(N,\delta^*_N)$, for fixed $N\in\mathcal{N}$.

The sequence $\{v^\alpha_N\}_{N<\infty}$ is a bounded sequence in $\ell^\infty(\mathbf{S},V)$. Consider any limit point of this sequence, say it is achieved along the subsequence $\{v^\alpha_{N_k}\}_{k=1,\ldots}$. By sequential compactness of $\mathcal{N}\times\mathcal{D}$, we have that $(\delta^*_{N_k})_k$ has a convergent subsequence that we denote by $(\delta^*_{N_k})_{N_k\in\mathcal{N}}$ again, with limit $\delta^*$ say.

Since the assertions of Theorem 3.4.1 hold for $\{X(N,\delta)\}_{(N,\delta)\in\mathcal{N}\times\mathcal{D}}$, this implies that $(N,\delta)\mapsto v^\alpha(N,\delta)$ is continuous on $\mathcal{N}\times\mathcal{D}$. In particular, we have

$$\lim_{k\to\infty} v^\alpha_{N_k} = \lim_{k\to\infty} v^\alpha(N_k,\delta^*_{N_k}) = v^\alpha(\infty,\delta^*) = v^\alpha_\infty(\delta^*).$$

Continuity of the map $(N,\delta)\mapsto v^\alpha(N,\delta)$, the fact that $v^\alpha(N_k,\delta^*_{N_k})$ solves the optimality equation for the $N_k$-perturbation by Theorem 3.4.2, and the continuity result of Theorem 3.3.3 together imply that $v^\alpha(\delta^*)$ solves the optimality equation for the $\infty$-perturbation, in other words, for the original MDP. Hence, $v^\alpha(\delta^*)=v^\alpha$ and $\delta^*$ is optimal. This holds for any limit point of $\{v^\alpha_N\}_N$. Since the solution of the optimality equation is unique, any limit point is equal to $v^\alpha$ and corresponding limit points of $\{\delta_N\}_N$ are optimal. This proves (i). For the proof of (ii), we consider a limit point of the sequence of policies $\{\delta_N\}_N$ (for any sequence of optimal policies for the $N$-perturbation, $N=1,2,\ldots$). Then choose a subsequence along which $\{v^\alpha_N\}_N$ converges and we apply the same argument as in the above. □

*The approach of extended parametrisation*

1. Start with a parametrised process $\{X(\delta)\}_{\delta\in\mathcal{D}}$, the original MDP. Our interest is in the structural properties of $v^\alpha$ and $\delta^*$. The assumptions of Theorem 3.4.2 must hold for this parametrised process.

2. Add a perturbation, parametrised by $N \in \mathcal{N}$. In this way we obtain an extended parametrised process $\{X(N,\delta)\}_{(N,\delta)\in\mathcal{N}\times\mathcal{D}}$. The extended parametrised process does not need to satisfy the product property of Assumption 3.4.2. However, all other assumptions from Theorem 3.4.2 are assumed to be satisfied for the extended parametrised process.

3. Fixing $N \in \mathcal{N}$, the parametrised process $\{X(N,\delta)\}_{\delta\in\mathcal{D}}$ is a perturbed process. It satisfies the product property of Assumption 3.4.2, and so all assumptions of Theorem 3.4.2 hold. Hence, there exists a unique solution $v_N^\alpha$ satisfying Eq. (3.7) and any maximiser $\delta_N^*$ is optimal.

4. If the perturbed process is uniformisable for all $N < \infty$, we can determine structural properties of $(v_N^\alpha, \delta_N^*)$ for all $N < \infty$ by e.g. value iteration.

5. Now Theorem 3.5.1 implies that $\lim_{N\to\infty} v_N^\alpha = v^\alpha$ and that any limit point of $(\delta_N^*)_N$ is optimal for the original model. As a conclusion, both the optimal policy and the minimum expected $\alpha$-discounted cost of the original model can be approximated by the corresponding quantities for the perturbed model, for large perturbation parameters.

*Remark* 3.5.2. Theorem 3.5.1 is strongly related to [52, Theorem 3.1]. The paper gives conditions for convergence of finite state MDP to infinite state processes. However the drift conditions imposed are more restrictive (cf. [66, Example 5.4]). In particular, the authors impose three extra conditions on the rate matrix, namely that $V$ is a moment function. Secondly, that

$$\sup_\delta q_x(N,\delta) \le V(x), \text{ for all } N \in \mathcal{N}.$$

Thirdly, they require a particular $V$-moment function $W$, namely $W = V^2$.

The last part of the chapter is an illustration of the application of the approach to a server farm model.

## 3.6 Optimal control of a server farm

Consider the server farm model studied by [1]. This model has an infinite server pool, implying that the transition rates are not bounded. To derive structural properties of the optimal policy the authors bound the departure rate. After uniformisation, analysis of the equivalent discrete time chain shows that a specific switching curve is optimal for the bounded rate model. However, this paper does not give any results on the original unbounded model.

We will demonstrate here that the same structural results apply for the unbounded model by using the approach of extended parametrisation.

The mathematical set-up is as follows. There is a Poisson stream of arrivals with rate $\lambda$. Each customer requires an exponential service time with parameter $\mu$. There is an infinite server pool, where servers can be in three states. They can be either active (on), turned off (off) or in standby modus (idle). After service completion the controller has two options, either turn the server off, or leave the server idle. A server in the idle state costs $c$ per unit time, due to energy consumption. Upon customer arrival, there are two possibilities.

**i)** There is an available idle server. Then the arriving customer is assigned one of these, and the server changes from idle to on.

**ii)** There are no idle servers. Then an off-server is turned on, to which the arriving customer is assigned, and instantaneous start-up costs $K$ have to be paid.

The goal is to minimise the total expected discounted cost over all stationary policies.

We will model this as follows. Let $i$ the number of idle servers and $j$ the number of busy servers. The state space $\mathbf{S}$ is given by

$$\mathbf{S} = \{(i,j) | i,j \in \mathbf{Z}_+\}.$$

Possible actions at service completion are either to turn the server off (0), or leave the server idle (1). The action space is

$$\mathcal{D}_{(i,j)} = \{0,1\}, \quad \text{for } (i,j) \in \mathbf{S}.$$

Hence, the set of stationary deterministic policies is $\mathcal{D} = \{0,1\}^{\mathbf{S}}$. Then the rate matrix $Q(\delta)$ is given by

$$q_{(i,j),(i',j')}(\delta) = \begin{cases} j\mu, & (i',j') = (i,j-1), \delta(i,j) = 0, \\ & \quad \text{or } (i',j') = (i+1,j-1), \delta(i,j) = 1, \\ \lambda, & (i',j') = (i-1,j+1), i > 0, \\ & \quad \text{or } (i',j') = (i,j+1), i = 0, \\ -(j\mu + \lambda), & (i',j') = (i,j). \end{cases}$$

The associated cost function $c(\delta)$ is given by

$$c_{(i,j)}(\delta) = ci + \lambda K \cdot \mathbb{1}_{\{i=0\}}, \quad (i,j) \in \mathbf{S}.$$

Notice that we have remodelled the instantaneous costs as a cost rate. This can be done without loss of generality.

As pointed out in the above, the rates $q_{(i,j)}(\delta) = j\mu + \lambda$ are not uniformly bounded as a function of state. To analyse this system, [1] assumes that the service rates are a concave, nondecreasing, bounded function $\mu(j)$ of the number of busy servers $j$ and thereby they make it uniformisable.

We will use this to define a suitable perturbation of the model, i.e. a uniformisable MDP, with the service rates a concave, nondecreasing and bounded function of the number of busy servers. As the original MDP is denoted by $\{X(\delta)\}_{\delta\in\mathcal{D}}$, we define a collection of perturbed MDPs $\{X(N,\delta\}_{N\in\mathcal{N},\delta,\in\mathcal{D}}$, with $\mathcal{N} = \{1,\ldots,\infty\}$. Let the rate matrix $Q(N,\delta)$ be given by

$$q_{(i,j),(i',j')}(N,\delta) = \begin{cases} (j\wedge N)\mu, & (i',j') = (i,j-1), \delta(i,j) = 0, \\ & \text{or } (i',j') = (i+1,j-1), \delta(i,j) = 1, \\ \lambda, & (i',j') = (i-1,j+1), i > 0, \\ & \text{or } (i',j') = (i,j+1), i = 0, \\ -((j\wedge N)\mu + \lambda), & (i',j') = (i,j). \end{cases}$$

The cost function remains unchanged.

Note that $X(\infty,\delta)$ coincides with the original unbounded model. On the other hand for each $N<\infty$, the $N$-perturbation is uniformisable and satisfies the service rate conditions of [1]. Hence, the structural properties of the value function $v_N^\alpha$ can be derived by value iteration. By virtue of the results in [1] it follows that the optimal policy for the $N$-perturbation, $N<\infty$, has the switching curve structure shown in Table 3.1. For a definition of the properties we refer to Section 7.3.

Table 3.1: *If it is optimal to turn the server off (respectively leave the server idle) in state $(i,j)$ then it is also optimal in the following states.*

| | leave idle | turn off | structural property of $v_N^\alpha$ |
|---|---|---|---|
| 1) | $\downarrow$: $(i,j-1)$ | $\uparrow$: $(i,j+1)$ | $\mathcal{S}uper(1,2)$ |
| 2) | $\nwarrow$: $(i-1,j+1)$ | $\searrow$: $(i+1,j-1)$ | $\mathcal{S}uper\mathcal{C}(1,2)$ |

With the approach of 'extended parametrisation', we are able to extend this result to the original unbounded model. The only thing remaining is to check that the assumptions of Theorem 3.5.1 hold. If the conditions hold, by virtue of the theorem we may conclude that a switching curve policy with the structure given in Table 3.1 is optimal for the original unbounded MDP. This yields the following result.

**Theorem 3.6.1.** *For the server farm model $\{X(\delta)\}_\delta$ there exists a deterministic policy with the threshold structure described in Table 1, that is $\alpha$-discount optimal within the class $\mathcal{S}$ of stationary policies.*

*Proof.* Note that the assumptions are of such nature that if they are satisfied by the extended parametrised process they are also satisfied by the parametrised process. As has been pointed out we have to verify the assumptions of Theorems 3.4.1 and 3.4.2. We will do so in a systematic way.

- It is clear that Assumption 3.2.1 holds for both the parametrised as the extended parametrised process, since there are no instantaneous jumps and the rate matrix is conservative.
- For Assumption 3.3.2 there are three properties to check.
  **i)** Continuity of $\delta \mapsto q_{(i,j),(i',j')}(N,\delta)$ for fixed $N \in \mathcal{N}$ is clear. Also, we have $\lim_{N\to\infty} q_{(i,j),(i',j')}(N,\delta) = q_{(i,j),(i',j')}(\infty,\delta)$ (for large $N$ these values are equal, for any fixed pair of states). As a consequence, it follows that $(N,\delta) \mapsto q_{(i,j),(i',j')}(N,\delta)$ is continuous on $\mathcal{N} \times \mathcal{D}$.
  **ii)** Let $0 < \gamma < \alpha$. Take $V(i,j) = \exp\{\epsilon(i+j)\}$, with

$$\epsilon = \frac{1}{2}\log(\gamma/\lambda + 1) > 0.$$

  Then $V$ clearly is a moment function. Moreover it is an $(\mathcal{N}\times\mathcal{D},\gamma)$-drift function. Indeed,

$$\begin{aligned}
&\sum_{(i',j')} q_{(i,j),(i',j')}(N,\delta)V(i,j) \\
&= e^{\epsilon(i+j)} \begin{cases} \min\{j,N\}\mu(e^{-\epsilon}-1)+\lambda(e^{\epsilon}-1), & \delta(i,j)=0,\ i=0, \\ \lambda(e^{\epsilon}-1), & \delta(i,j)=1,\ i=0, \\ \min\{j,N\}\mu(e^{-\epsilon}-1), & \delta(i,j)=0,\ i>0, \\ 0, & \delta(i,j)=1,\ i>0, \end{cases} \\
&\leq \lambda(e^{\epsilon}-1)e^{\epsilon(i+j)} \\
&\leq \lambda(e^{2\epsilon}-1)e^{\epsilon(i+j)} \\
&= \gamma e^{\epsilon(i+j)} = \gamma V(i,j).
\end{aligned}$$

Thus, $V$ is an $(\mathcal{N}\times\mathcal{D},\gamma)$ drift function for $X(N,\delta)$, uniformly on $\mathcal{N}\times\mathcal{D}$.

**iii)** Take $W(i,j)=\exp\{2\epsilon(i+j)\}$, then $W/V=V$ is a moment function. Hence, $W$ is a $V$-moment function, in particular, $W$ is a $(\mathcal{N}\times\mathcal{D},\theta)$-drift $V$-moment function for $\theta=\lambda(e^{4\epsilon}-1)$, since

$$\begin{aligned}
\sum_{(i',j')} q_{(i,j),(i',j')}(N,\delta)W(i,j) &\leq \lambda(e^{4\epsilon}-1)e^{2\epsilon(i+j)} \\
&= \theta e^{2\epsilon(i+j)} = \theta W(i,j).
\end{aligned}$$

- Consider Assumption 3.4.1.

**i)** $(N,\delta)\mapsto c_{(i,j)}(N,\delta)$ is clearly continuous on $\mathcal{N}\times\mathcal{D}$, for any $(i,j)\in S$.

**ii)** Take $c_V=\frac{c}{\lambda\epsilon}+\lambda K$, then for any $(i,j),(N,\delta)$

$$\begin{aligned}
\frac{|c_{(i,j)}(n,\delta)|}{V(i,j)} &= \frac{ci/(\lambda+j\mu)+\lambda K\mathbf{1}_{\{i=0\}}}{\exp\{\epsilon(i+j)\}} \\
&\leq \frac{(c/\lambda)i+\lambda K}{1+\epsilon i} \\
&\leq \frac{c}{\lambda\epsilon}+\lambda K = c_V.
\end{aligned}$$

Hence the supremum over all $(i,j),(N,\delta)$ is also bounded by $c_V$.

- Condition (i) of Assumption 3.4.2 holds for both the parametrised process and the extended parametrised process.

**i)** The parameter set is a product space $\mathcal{D} = \prod_{(i,j)\in\mathbf{S}} \mathcal{D}_{(i,j)}$, with $D_{(i,j)}$ a finite set, hence compact and metric for each state $(i,j) \in \mathbf{S}$. The set $\mathcal{N}$ is compact; hence, $\mathcal{N} \times \mathcal{D}$ is compact.

Condition (ii) of Assumption 3.4.2 only holds for the parametrised process $\{X(N,\delta)\}_\delta$, for $N \in \mathcal{N}$ fixed, and not for the extended parametrised process.

**ii)** $\{Q(\delta)\}_{\delta\in\mathcal{D}}$ and $\{c(\delta)\}_{\delta\in\mathcal{D}}$ both have the product property. In other words, the transition rates and the cost rates in state $(i,j)$ depend only on the action in state $(i,j)$.

$\square$

# 4 Power control of a server farm

This chapter is based on Blok et al. [17], in preparation.

## 4.1 Introduction

In recent years, awareness towards energy consumption in server farms and data centres has become more and more important. Consequently, optimal design of server farms has become an active area of research. One aspect in server farm management that has drawn special attention, is how to decrease the cost for servers that are unnecessarily idle. Server farms exhibit an intrinsic trade-off between the heavy energy consumption in case of too many idle servers and set-up or delay costs in case of lack of available idle servers.

We consider a server farm with ample service capacity. Servers that are not active can be either idle or off. Idle servers consume energy, which is modelled as a holding cost per idle server and per unit time. Many models include set up times for switching on off servers (cf. [32], [33] and [48]). Instead, we follow [1] that models this phenomenon by instantanuous start up costs. We also allow to incorparate nonnegative costs for switching off a server upon finishing service. Control is exercised at a moment of a service completion, where the decision is made to turn a server off or leave it idle at the moment it finishes service. This can be modelled as a continuous time Markov decision process (MDP) with unbounded jump rates, with the instantaneous switch on and switch off cost remodelled as cost rates.

We are interested in the policy that minimises the expected $\alpha$-discounted cost and average cost. In particular, our aim is to derive the structure of an optimal policy.

This model was introduced and studied by Adan et al. [1]. In their paper they show that a certain switching curve policy is optimal. However the results are restricted to the case where the total server capacity is finite. This is due to the fact that the standard techniques to analyse structural properties of MDPs are limited to bounded jump rate MDPs. Similar monotonicity results have been obtained in [48], however a good comparison is difficult due to different modelling choices.

This problem with unbounded jumps touches upon a gap in the literature, for which so far no systematic solution has been given. MDPs with unbounded rates are not uniformisable. This hampers applicability of the value iteration (VI) algorithm, as a method for deriving structural properties of optimal policies. Thus, a truncation should be used to make the MDP uniformisable. Unfortunately, in the average cost criterion, this causes two problems. First, due to the existence of transient policies in the server farm model under consideration, the VI algorithm is not guaranteed to converge under the average cost criterion. Secondly, one needs a continuity argument to transfer the results from the truncated MDPs to the original MDP. To the best of our knowledge no such theory exists that covers the present model. However, under the total discounted cost criterion, convergence of VI holds under very mild conditions. Hence for the bounded rate $\alpha$-discounted cost problem structural properties can be shown via VI. Moreover, these results can be transferred to unbounded rate $\alpha$-discounted MDPs via the limit theorem on parametrised Markov processes from Chapter 3. This has been already shown in that chapter in case of zero switch off cost.

In this chapter we extend this result also to the case of non-zero switch off cost, and thus we show that also in this case there exists a switching curve $\alpha$-discount optimal policy. Via the continuous time vanishing discount approach of Chapter 2, switching curve optimality applies as well for the average cost criterion.

A second focus of the chapter is on bounds for the switching curve. As a function of the discount factor $\alpha$, we determine bounds on the optimal switching curve. In the limit of discount factor to 0, the bounds converge to a bound on the switching curve for the average optimality criterion, due to applicability of the vanishing discount principle. This then can be exploited to show (strong) Blackwell optimality of a switching curve policy. It can also be exploited for a finite algorithm to compute an optimal policy, for both the $\alpha$-discounted cost and average cost criterion.

We introduce the mathematical model in Section 4.2.1. We then define the various optimality criteria of interest and summarise our main results in Section 4.2.2.

In Section 4.3, we prove switching curve optimality for both the $\alpha$-discounted and the average cost criteria. The section is supported by Section 4.5, which consists of the propagation results in an event based dynamic programming framework. The propagation results are similar to those in [1], but extended to allow for non-zero switch off cost. It is not necessary to use the more refined smoothed rate truncation method, that has proven to be succesful for obtaining structural properties in similar models (cf. [11], [16] and Chapters 5 and 6).

This server farm model was also examined in Kappetein [42]. Coupling techniques are used there to show optimality of a switching curve policy for the average cost criterion. The ideas of this work were applied to the $\alpha$-discounted cost criterion in Van der Velde [73]. Moreover[42] and [73] localise for which problems extreme policies are optimal and for which it is a true switching curve. However, these works still used an instantaneous switch on and switch off cost model, instead of cost rates, as we do here. This complicates the derivations. Moreover, rigorous proofs lack in these works. Section 4.4 remedies this and provides the coupling proofs in a more precise manner, supported by Section 4.6.

This chapter can be seen as an example to the problem of deriving properties for unbouded rate average cost MPDs. The method goes from truncated discounted MDPs via unbounded jump discounted cost MDPs to the unbounded jump average cost MDP. It remains a burning question whether it is possible to take a directer approach, using truncated average cost MDPs. We would then need a result regarding convergence of average cost VI that does not rely on strong ergodicity assumptions. Moreover, we would need sufficient conditions validating the continuity of performance measures of the truncated MDPs as they approach the original Markov decision process, which so far seem to be lacking.

## 4.2 Model and main results

### 4.2.1 Model description

The formal description of the power controlled server farm model is as follows. Customers arrive according to a Poisson ($\lambda$) process. Each customer requires an exponentially distributed amount of service with parameter $\mu > 0$. For design reasons the server pool is assumed to be unbounded. Servers can be in three states: active, turned off or in standby modus. They will be referred to as 'busy', 'off' and 'idle'.

After a service completion the controller has to decide to either turn the server off incurring an instantaneous cost $K^{\mathrm{off}}$, or to leave the server idle. Each idle server consumes power, which is assumed to cost $c > 0$ per unit time.

Upon customer arrival, there are two possibilities. If an idle server is available, the arriving customer is assigned one of these, and the server changes from idle to on, instantaneously. If no idle servers are available, an off server is instantaneously turned on, and an instantaneous start-up cost $K^{\mathrm{on}}$ is incurred. The goal is to minimise the *expected total $\alpha$-discounted* and the *expected average cost*.

This control problem can be modelled as a continuous time Markov decision process on the state space $\mathbf{S}$, with

$$\mathbf{S} = \mathbf{Z}_+^2,$$

where state $(x_1, x_2)$ corresponds to the presence of $x_1$ busy servers and $x_2$ idle ones. Note that it is not necessary to keep track of the number of off servers.

The possible control actions at a service completion are either to turn the server off (OFF), or leave the server idle (IDLE). A service completion can only happen in the states of $\mathbf{S}^* = \{x \in \mathbf{S}, x_2 \geq 1\}$. Hence, for $x \in \mathbf{S}^*$ the decision or action set is $\mathcal{D}_x = \{\text{IDLE}, \text{OFF}\}$. For $x \in \mathbf{S} \setminus \mathbf{S}^*$ only one action is available, say 0, thus $\mathcal{D}_x = \{0\}$, $x \in \mathbf{S} \setminus \mathbf{S}^*$. As a consequence, the collection of stationary, deterministic policies is

$$\mathcal{D} = \prod_{x \in \mathbf{S}} \mathcal{D}_x.$$

We denote a stationary, deterministic policy by $\delta$, and by $\delta_x$ the corresponding decision in state $x$.

Then, the rate matrix $Q(\delta)$, $\delta \in \mathcal{D}$, is given by

$$q_{xy}(\delta) = \begin{cases} x_2\mu & y = x - e_2, \delta_x = \text{OFF or } y = x + e_1 - e_2, \delta_x = \text{IDLE}; \\ \lambda & y = x - e_1 + e_2, x_1 > 0, \text{ or } y = x + e_2, x_1 = 0; \\ -(x_2\mu + \lambda) & y = x, \end{cases}$$

for $x, y \in \mathbf{S}$. The associated cost function $c : \mathbf{S} \times \mathcal{D} \to \mathbb{R}$ is given by

$$c_x(\delta) = cx_1 + \lambda K^{\mathrm{on}} \cdot \mathbb{1}_{\{x_1 = 0\}} + x_2\mu K^{\mathrm{off}} \cdot \mathbb{1}_{\{\delta_x = \text{OFF}\}}.$$

Notice that we have remodelled the instantaneous costs as a cost rate, incurred before the moment the actual instantaneous cost is paid. This will turn out to be a complicating factor in the derivation of bounds in Section 4.4.

One may also randomise between the actions in a state. In state $x$ this yields a collection of probability distributions, $\Pi_x$ say, over $\mathcal{D}_x$. Then, $\Pi = \prod_{x \in \mathbf{S}} \Pi_x$ denotes the collection of stationary policies, and $\pi = \{\pi_x\}_{x \in \mathbf{S}}$ denotes a generic element. Thus, $Q(\pi)$ is the $q$-matrix with elements

$$q_{xy}(\pi_x) = \pi_x(\text{IDLE}) q_{xy}(\text{IDLE}) + \pi_x(\text{OFF}) q_{xy}(\text{OFF}), \quad x, y, x_2 \neq 0 \in \mathbf{S},$$

and $c(\pi)$ is the corresponding cost rate vector with elements

$$c_x(\pi_x) = \pi_x(\text{IDLE})c_x(\text{IDLE}) + \pi_x(\text{OFF})c_x(\text{OFF}), \quad x \in \mathbf{S}.$$

Each stationary policy $\pi \in \Pi$ defines a probability distribution on an underlying measurable space $(\Omega, \mathcal{F})$, where $\Omega$ is in fact an infinite sequence of potential states, decisions, and sojourn times per state, cf. [44, 51]. On this space one can define a stochastic process $(X, D)$, with

$$\begin{aligned}(X, D) : \Omega \quad \mapsto \quad & \{f : [0, \infty) \to \mathbf{S} \mid f \text{ right-continuous}\} \\ & \times \{g : [0, \infty) \to \{\text{IDLE}, \text{OFF}\} \mid g \text{ right-continuous}\},\end{aligned}$$

a filtration $\{\mathcal{F}_t\}_t \subset \mathcal{F}$, to which $(X, D)$ is adapted, and a probability distribution $\mathsf{P}^\pi_\nu$ on $(\Omega, \mathcal{F})$, such that $X$ is the minimal Markov process with $q$-matrix $Q(\pi)$, for each initial distribution $\nu$ on $\mathbf{S}$. Denote by $X_t = (X_{t,1}, X_{t,2})$, and $D_t$ the state and decision at time $t$ respectively. Denote by $P(\pi) = \{\, p_{xy}(\pi)\}_{x,y \in \mathbf{S}}$ the corresponding transition probability matrix, and by $\mathsf{E}^\pi_\nu$ the corresponding expectation operator. Furthermore,

$$\mathsf{P}^\pi_\nu\{X_t = y, D_t = d\} = \sum_{x \in \mathbf{S}} \nu_x \, p_{t,xy}(\pi)\pi_y(d).$$

We note that $X$ is also a standard and stable Markov process under $\mathsf{P}^\pi_\nu$. Occasionally it may be convenient to use the notation $X(\pi)$ to denote the Markov process with distribution $\mathsf{P}^\pi$.

## 4.2.2 Optimality Criteria and Main results

The aim is to study the structure of an optimal policy for the following criteria. Note that we restrict to optimality within the class $\Pi$ of stationary policies. This can be extended to more general policy classes (cf. [35, 44, 51]). First we give the definitions of the optimality criteria that we consider. The proof of Theorem 4.2.1 provides the arguments showing that the server farm model satisfies conditions guaranteeing these criteria to be well-defined and optimal policies to exist within the class $\mathcal{D}$.

For $\alpha > 0$, the *expected total $\alpha$-discounted cost $v^\alpha(\pi)$ associated with policy* $\pi \in \Pi$ is given by

$$v^\alpha_x(\pi) = \mathsf{E}^\pi_x \int_0^\infty e^{\alpha t} c_{X_t}(D_t) dt, \quad x \in \mathbf{S}.$$

The *minimum expected total $\alpha$-discounted cost value function $v^\alpha$* is defined by

$$v^\alpha_x = \inf_{\pi \in \Pi} v^\alpha_x(\pi).$$

If for some $\pi \in \Pi$ it holds that $v^\alpha(\pi) = v^\alpha$, then $\pi$ is said to be *$\alpha$-discount optimal.*

The *expected average cost* $\mathsf{g}(\pi)$ *associated with policy* $\pi \in \Pi$ is given by

$$\mathsf{g}_x(\pi) = \limsup_{T\to\infty} \frac{1}{T}\mathsf{E}_x^\pi \int_0^T c_{X_t}(D_t)dt.$$

The *minimum expected average cost* $\mathsf{g}$ is given by

$$\mathsf{g}_x = \inf_{\pi\in\Pi} \mathsf{g}_x(\pi).$$

If for some $\pi \in \Pi$ it holds that $\mathsf{g}(\pi) = \mathsf{g}$, then $\pi$ is said to be *average optimal.* Suppose policy $\pi^*$ is $\alpha$-discounted optimal for all $\alpha \in (0, \alpha_0)$ for some $\alpha_0 > 0$. Then $\pi^*$ is said to be *strong Blackwell optimal.* This notion has been introduced by Blackwell [15].

In this chapter we will prove the following two main results. The first one concerns average and $\alpha$-discount optimality of a switching curve policy. The second provides an explicit bound in terms of the input parameters of the process on the region where the decision to idle a server finishing service can be optimal. Denote by $e_i$ the $i$-th unit vector.

**Definition 4.2.1.** A deterministic policy $\pi \in \Pi$ is called SC-policy (Switching Curve-policy), if

**i)** $\pi_x = \text{OFF} \implies \pi_{x+e_2}, \pi_{x+e_1-e_2} = \text{OFF};$

**ii)** $\pi_x = \text{IDLE} \implies \pi_{x-e_2}, \pi_{x-e_1+e_2} = \text{IDLE},$

provided the resulting states belong to $\mathbf{S}^*$.

**Theorem 4.2.1.** *i) There exists an* SC*-policy that is $\alpha$-discount optimal policy within the class* $\Pi$*.*

*ii) There exists an* SC*-policy that is average optimal within the class* $\Pi$.

Next, define $B = \{x \in \mathbf{S}^* \mid x_1 = 0\}$, which is the $x_2$-axis without the origin. Further $\Delta(n) = \{x \in \mathbf{S} \mid x_1 + x_2 \le n\}$. If $c > \alpha K^{\text{off}}$, define for $\alpha \ge 0$

$$n_1(\alpha) = \lambda\frac{K^{\text{off}} + K^{\text{on}}}{c - \alpha K^{\text{off}}},$$

and

$$\begin{aligned} n_0(\alpha) &= \max\Big\{4, 2 + k(\rho), \\ &\quad \max\{n_1(\alpha), 2\} + \frac{(\alpha+\mu+\lambda)\eta(\alpha)k^2(\rho)}{\mu(c - \alpha K^{\text{off}})}\Big[K^{\text{off}}(\lambda + 2\mu) + K^{\text{on}} \cdot 2\lambda\Big]\Big\}\end{aligned}$$

where

$$\eta(\alpha) = \Big(\frac{\alpha+\mu+\lambda}{\mu}\Big)^{n_1(\alpha)}, \quad k(\rho) = \frac{(1+\rho)^2}{\rho}e^{\rho}, \quad \text{and } \rho = \frac{\lambda}{\mu}. \tag{4.1}$$

**Theorem 4.2.2.** *1. Consider the $\alpha$-discount optimality criterion.*

**i)** *If $c > (\lambda+\alpha)K^{\mathrm{off}} + \lambda K^{\mathrm{on}}$, then the policy $\delta$ with $\delta_x = \text{OFF}$, $x \in \mathbf{S}^*$, is $\alpha$-discount optimal.*

**ii)** *If $\alpha K^{\mathrm{off}} < c \leq (\lambda+\alpha)K^{\mathrm{off}} + \lambda K^{\mathrm{on}}$, then there exists an $\alpha$-discount optimal policy $\delta$ with the property that*

**a)** *$\delta_x = \text{IDLE}$, $x \in B$;*

**b)** *$\delta_x = \text{OFF}$, $x \in \mathbf{S}^* \setminus \{\Delta(n_0(\alpha)) \cup B\}$.*

**iii)** *If $c \leq \alpha K^{\mathrm{off}}$, then $\delta$ with $\delta_x = \text{IDLE}$, $x \in \mathbf{S}^*$, is $\alpha$-discount optimal.*

*2. Consider the average optimality criterion.*

**i)** *If $c > \lambda(K^{\mathrm{off}} + K^{\mathrm{on}})$, then the policy $\delta$ with $\delta_x = \text{OFF}$, $x \in \mathbf{S}^*$, is average optimal.*

**ii)** *If $0 < c \leq \lambda(K^{\mathrm{off}} + K^{\mathrm{on}})$, then there exists an average optimal policy with the property that*

**a)** *$\delta_x = \text{IDLE}$, $x \in B$;*

**b)** *$\delta_x = \text{OFF}$, $x \in \mathbf{S}^* \setminus \{\Delta(n_0(0)) \cup B\}$.*

*3. Denote by $\delta^\alpha$ an $\alpha$-discount* SC *optimal policy of the type described under 1. Then any limit point of the sequence $\delta^\alpha$, $\alpha \downarrow 0$, is an* SC *strong Blackwell optimal policy.*

The SC-policy is illustrated in Figure 4.1. The red states form a positive recurrent class in the Markov process associated with the stationary, deterministic SC policy. The policy that idles a busy server upon finishing service in the grey area states, and switches off in the white area states.

The proof of the above theorem hinges on two very crucial properties of the server farm model. The first one is that the process associated with the number of customers in the system is an M/M/$\infty$-queue. This is precisely the process $\{X_{t,2}\}_{t\geq 0}$ associated with the number of busy servers. Thus, the stationary number of busy servers has a Poisson distribution with parameter $\lambda/\mu$.

The second one is the cleverly exploited property in [42], that any triangle $\Delta(n)$ can only be left at state $(0,n)$, i.e. it is an *exit state* of $\Delta(n)$ (cf. [43]).

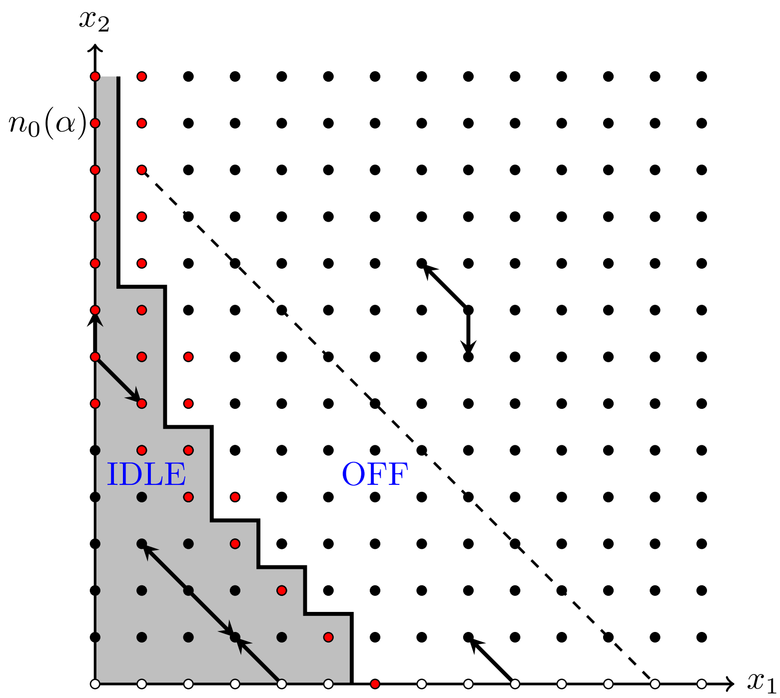



Table 4.1: The optimal policy has a switching curve structure

The combination of the two properties yields that the time to reach $(0, n+1)$ from any state $(i, j) \in \Delta(n)$ is independent of the policy, and only depends on the value $j$. In fact, that it is equal (in distribution) to the time to reach state $n+1$ from state $j \leq n$ in the M/M/$\infty$-queue. This yields the property formulated in Lemma 4.2.3. First we need to introduce some notation.

***n*-restricted MDP** The $n$-restriction is the following MDP with a finite state space $\Delta(n)$, and finite decision spaces. The decision spaces, and the transition and cost rates on $\Delta(n)$ are identical to the corresponding descriptors of the server farm model, however, in state $(0, n)$ only one decision is available, that is, $\mathcal{D}_{(0,n)} = \{0\}$. Additionally, $(0, n)$ is an absorbing zero cost state.

Some specific notation: $A(n) = \{x \in \mathbf{S} \mid x_1 + x_2 = n, x_1 \neq 0\}$. Furthermore, $v^\alpha(n) = \{v^\alpha_x(n)\}_{x \in \Delta(n)}$ is the $\alpha$-discounted value function associated with the $n$-restriction, $\alpha \geq 0$. This is finite, also for $\alpha = 0$, since the $n$-restriction defines a uniformisable finite state, finite decision MDP (cf. [26, 64]), with absorbing zero cost state $(0, n)$.

We will also need to introduce some quantities associated with the M/M/$\infty$-queue. The states are indexed by $i \in \{0, 1, \dots\}$. Let $\tau_i$ be the hitting time of state $i$. Denote $t_i = \mathsf{E}_i e^{-\alpha\tau_{i+1}}$.

Then, it is simply checked that the following recursion holds (cf. also [28, Eq. (2) and above]): $t_0 = \lambda/(\lambda+\alpha)$ and

$$t_l = \frac{\lambda}{\lambda + \alpha + (1 - t_{l-1})l\mu}, \quad l \geq 1.$$

Furthermore, for $l < i$ it holds that $\mathsf{E}_l e^{-\alpha\tau_{i+1}} = t_l \cdots t_i$.

**Lemma 4.2.3.** *For $x \in \Delta(n) \setminus \{(0,n)\}$ it holds that*

$$v_x^\alpha = v_x^\alpha(n) + \prod_{l=x_2}^{n-1} t_l \cdot v_{(0,n)}^\alpha.$$

*For the $\alpha$-discounted cost criterion the following holds. If $\pi \in \Pi$ is optimal for the server farm model, then $\{\pi_x\}_{x\in\Delta(n)\setminus\{0,n)\}}$ is optimal for the $n$-restriction. Vice versa, if $\{\pi_x\}_{x\in\Delta(n)\setminus\{(0,n)\}}$ is optimal for the $n$-restriction, then $\{\pi_x\}_{x\in\Delta(n)\setminus\{(0,n)\}}$ can be extended to a stationary optimal policy for the server farm model. The latter implication holds as well for the average cost criterion.*

We would like to point out that not every average cost optimal policy needs to be a solution to the average cost optimality equation Eq. (4.10). This is responsible for the asymmetry in the above lemma with respect to the two optimality criteria.

Theorem 4.2.1, Theorem 4.2.2 and Lemma 4.2.3 provide an efficient computation of an ($\alpha$-discounted, or average) optimal policy. We give the algorithm for the $\alpha$-discounted cost. The corresponding algorithm for the average cost can obtained from the $\alpha$-discounted cost algorithm, by inserting the value $\alpha = 0$.

Note that the algorithm stops after at most $n_0(\alpha)$ iterations of step 1, with an SC optimal policy. This is an advantage above algorithms as value iteration, that yield only approximations. The validity of this algorithm is proven in Lemma 4.4.4.

## 4.3 Roadmap and proofs

**$\alpha$-discount optimality** First of all we will consider the $\alpha$-discounted cost criterion. The paper [1] has proven $\alpha$-discount optimality of an SC-policy, for the case $K^{\mathrm{off}} = 0$, if the service rates are truncated at maximum value, say $N\mu$

**Algorithm 3** Computation of an SC optimal policy $\delta^\alpha$ for the $\alpha$-discounted criterion and of $\delta^0$ for the average cost criterion

**0)** Initialise: if $c > (\lambda+\alpha)K^{\mathrm{off}} + \lambda K^{\mathrm{on}}$, then put $\pi_x = \textsc{off}$, $x \in \mathbf{S}^*$, STOP;
if $c \le \alpha K^{\mathrm{off}}$, then put $\pi_x = \textsc{idle}$, $x \in \mathbf{S}^*$, STOP;
otherwise: put $\mathcal{D}^0_{(0,0)} = \{0\}$, $v^\alpha_{(0,0)}(0) = 0$; $\pi^0_{(0,0)} = 0$, $n = 1$, goto step 1.

**1)** Computation of an optimal policy on $\Delta(n) \setminus \{(0,n)\}$.

- For $x \in A(n) \cup \{(0,n-1)\}$ do:
  - put $\mathcal{D}^n_{(0,n-1)} = \{\textsc{idle}\}$;
  - if $\pi^{n-1}_{x-e_2} = \textsc{off}$, or $\pi^{n-1}_{x-e_1} = \textsc{off}$, put $\mathcal{D}^n_x = \{\textsc{off}\}$;
  - otherwise, put $\mathcal{D}^n_x = \mathcal{D}_x$.
- Put $v^\alpha_{(0,n)}(n) = 0$,
$$v^\alpha_{(0,n-1)}(n) = \frac{v^\alpha_{(1,n-2)}(n-1)\mu\mathbb{1}_{\{n>1\}} + \lambda K^{\mathrm{on}}}{\lambda + (n-1)\mu(1-t_{n-2})\mathbb{1}_{\{n>1\}} + \alpha};$$
and put
$$v^\alpha_x(n) = v^\alpha_x(n-1) + \prod_{l=x_2}^{n-2} t_k \cdot v^\alpha_{(0,n-1)}(n),\ x \in \Delta(n-1)\setminus\{(0,n-1)\}.$$
- Solve the following linear system in the unknowns $u_x$, $x \in A(n)$, with $u_{(0,n)} = 0$ (e.g. by policy iteration),
$$\begin{aligned}(\alpha+\lambda+x_2\mu)u_x &= c_x + \lambda u_{x-e_1+e_2}\mathbb{1}_{\{x_1>1\}}\\ &\quad + x_2\mu\cdot\min\Big\{\big(K^{\mathrm{off}} + v^\alpha_{x-e_2}(n)\big)\big(\mathbb{1}_{\{\textsc{off}\in\mathcal{D}^n_x\}} + \infty\mathbb{1}_{\{\textsc{off}\notin\mathcal{D}^n_x\}}\big),\\ &\qquad u_{x+e_1-e_2}\big(\mathbb{1}_{\{\textsc{idle}\in\mathcal{D}^n_x\}} + \infty\mathbb{1}_{\{\textsc{idle}\notin\mathcal{D}^n_x\}}\big)\Big\}. \qquad (4.2)\end{aligned}$$
- Put $\pi^n_x = \pi^{n-1}_x$, $x \in \Delta(n) \setminus A(n)$. Put $\pi^n_x$ equal to the minimising decision in Eq. (4.2); in case of a tie, select $\pi^n_x = \textsc{idle}$, $x \in A(n)$.

**2)** Stop, if $\pi^n_x = \textsc{off}$ for $x \in A(n)$. Then, $\delta^\alpha_x = \pi^n_x$, $x \in \Delta(n) \setminus \{(0,n)\}$. The values $\delta^\alpha_x$, $x \notin \Delta(n) \setminus \{(0,n)\}$ can be determined from Theorem 4.2.2. Otherwise, $n := n+1$, goto step 1.

(cf. Eq.(4.4) below). In Section 3.6 the case $K^{\text{off}} = 0$ is studied as well, and $\alpha$-discount optimality of an SC-policy is proved for the unbounded-rate MDP. The approach is a generic one validating to take the limit for the truncation in [1] to vanish. The case $K^{\text{off}} > 0$ is largely analogous to the proofs in [1] and in Chapter 3. We will not give all details, but we will describe in general terms how to tackle this problem.

First of all, we need $v^\alpha$ to be finite, as well as a solution to the $\alpha$-discount optimality equation, specifically

$$\begin{aligned}\alpha u_x &= cx_1 + \lambda\big(K^{\text{on}}\mathbb{1}_{\{x_1=0\}}u_{x+e_2} + u_{x-e_1+e_2}\mathbb{1}_{\{x_1>0\}}\big) - (\lambda + x_2\mu)u_x \\ &\qquad +x_2\mu \cdot \min\{K^{\text{off}} + u_{x-e_2}, u_{x-e_2+e_1}\}. \qquad (4.3)\end{aligned}$$

Furthermore, if $\delta$ is a deterministic policy that chooses the minimising actions in Eq. (4.3), then we need to validate that $v^\alpha(\delta) = v^\alpha$, so that $\delta$ is optimal over $\Pi$. In the present model this is true under the condition that there exists a policy with finite $\alpha$-discounted cost (cf. [70, 63]), since the problem is a cost minimisation problem with finite decision spaces and non-negative cost rates (negative dynamic programming). Further $v^\alpha$ is the minimum non-negative solution to Eq. (4.3).

In order to guarantee the existence of an SC $\alpha$-discount optimal policy, it is sufficient that $v^\alpha$ is supermodular and superconvex.

**Definition 4.3.1.** Let $v : \mathbf{S} \to \mathbb{R}$. Then $v$ is supermodular, if

$$v_{x+e_1+e_2} - v_{x+e_1} - v_{x+e_2} + v_x \geq 0, \text{ for } x \in \mathbf{S};$$

and $v$ is superconvex, if

$$v_{x+2e_1} - v_{x+e_1} - v_{x+e_1+e_2} + v_{x+e_2} \geq 0, \text{ for } x \in \mathbf{S}.$$

Denote $\mathcal{S}uper = \{v : \mathbf{S} \to \mathbb{R} \,|\, v \text{ supermodular}\}$, and $\mathcal{S}uper\mathcal{C} = \{v : \mathbf{S} \to \mathbb{R} \,|\, \text{v superconvex}\}$.

We show that supermodularity and superconvexity yield the desired result.

**Lemma 4.3.1.** *If $v^\alpha \in \mathcal{S}uper \cap \mathcal{S}uper\mathcal{C}$, then there exists a minimising (stationary) deterministic policy $\delta^\alpha$ in Eq. (4.3) that is an* SC*-policy. This policy $\delta^\alpha$ is thus $\alpha$-discount optimal.*

*Proof.* Let $\delta$ be a minimising policy in Eq. (4.3). Suppose that $\delta_x = \text{OFF}$, i.e. turning off is optimal in state $x$. Then Eq. (4.3) implies that $v^\alpha_{x+e_1-e_2} \geq v^\alpha_{x-e_2} + K^{\text{off}}$. Now $\mathcal{S}uper$ yields that $v^\alpha_{x+e_1} \geq v^\alpha_x + K^{\text{off}}$, implying that turning off in $x + e_2$ is optimal.

Furthermore, $\mathcal{S}uper\mathcal{C}$ yields that $v_{x+2e_1-2e_2} \geq v_{x+e_1-2e_2} + K^{\text{off}}$. Thus, turning off in $x + e_1 - e_2$ is optimal.

Similarly, suppose that $\delta_x = \textsc{idle}$. Then, $\mathcal{S}uper$ yields that idling in state $x - e_2$ is optimal, and $\mathcal{S}uper\mathcal{C}$ implies that idling in state $x - e_1 + e_2$ is optimal. $\square$

Note, that there may be ties, where both OFF and IDLE can be minimisers, hence optimal decisions. Thus, there may be optimal policies that are not an SC-policy.

In order to show the structural results from Theorem 4.2.1, value iteration (VI) is a proper method. This requires a uniformisation step and a subsequent time discretisation step. However, the first is hampered by the fact that the jump rates are unbounded as a function of state and decision pairs. Hence, one must perform a bounded jump rate perturbation that leaves structural properties invariant in order to apply VI.

Consequently, one needs continuity properties as a function of a vanishing perturbation parameter, so that structural properties of the original unperturbed model follow from the perturbed ones. Notice that, whereas the continuous time MDP involves a minimisation over the set of Markov processes $\{X(\pi)\}_{\pi\in\Pi}$, introducing a perturbation extends the set of Markov processes under consideration. Denoting the perturbation parameter by $N \in \mathcal{N}$, where $\mathcal{N}$ is the set of all perturbation parameters. Let one parameter be associated with the original unperturbed problem. The set of Markov processes to be considered is a parametrised one: $\{X(N, \pi)\}_{N\in\mathcal{N},\pi\in\Pi}$. Thus, the perturbations only affect the transition rates and not the cost rates. Then, $\{X(N, \pi\}_{\pi\in\Pi}$ defines an MDP as well, for any $N \in \mathcal{N}$.

The perturbation we will use, is the case studied in [1], where the service rates have been truncated. Then $\mathcal{N} = \{\mathbf{Z}_{>0}\} \cup \{\infty\}$, where parameter $N = \infty$ stands for the unperturbed MDP. Given $N \in \mathcal{N}$, the corresponding MDP has rates:

$$q_{xy}(N, \delta_x) = \begin{cases} \{x_2 \wedge N\}\mu & y = x - e_2, \delta_x = \textsc{off} \\ & \quad \text{or } y = x + e_1 - e_2, \delta_x = \textsc{idle}; \\ \lambda, & y = x - e_1 + e_2, x_1 > 0 \\ & \quad \text{or } y = x + e_2, x_1 = 0; \\ -(\{x_2 \wedge N\}\mu + \lambda) & y = x, \end{cases} \tag{4.4}$$

for $x \in \mathbf{S}$, $\delta_x \in \mathcal{D}_x$. We call this the $N$-perturbed MDP.

Again, in this particular case of a cost minimisation, the same weak conditions as mentioned above, ensure that for each $N$-perturbed MDP the $\alpha$-

discounted cost value function $v^\alpha(N)$ is finite and a solution to the corresponding $\alpha$-discount optimality equation

$$\begin{aligned}\alpha u_x &= x_1 c + \lambda\big\{\mathbb{1}_{\{x_1>0\}} u_{x-e_1+e_2} + \mathbb{1}_{\{x_1=0\}}(u_{x+e_2} + K^{\text{on}})\big\} \\ &\quad + \{x_2 \wedge N\}\mu \min\{U_{x+e_1-e_2}, u_{x-e_2} + K^{\text{off}}\} - (\lambda + \{x_2 \wedge N\}\mu)u_x. \quad (4.5)\end{aligned}$$

Furthermore, if $\delta$ is a deterministic policy that chooses minimising actions in Eq. (4.5), then $v^\alpha(N,\delta) = v^\alpha(N)$, and so $\delta$ is optimal over $\Pi$. However, we will also need that $v^\alpha(N) \to v^\alpha(\infty) = v^\alpha$, $N \to \infty$. This will follow from Theorem 4.3.2 below. Suitable conditions for this convergence have been derived in Chapter 3.

As a consequence, if $v^\alpha(N) \in \mathcal{S}uper\mathcal{C} \cap \mathcal{S}uper$, $N < \infty$, then $v^\alpha \in \mathcal{S}uper\mathcal{C} \cap \mathcal{S}uper$. So, we are only left to show that $v^\alpha(N) \in \mathcal{S}uper\mathcal{C} \cap \mathcal{S}uper$.

To do so, we will uniformise the $N$-perturbed MDP. It is convenient to first scale the rates, so that the maximum jump rate $\lambda + N\mu = 1$. Uniformisation yields a discrete time MDP (cf. [64] and Chapter 2) with transition matrix $P(\delta) = I + Q(\delta)$, direct cost $\bar{c}(\delta) = c(\delta)/(1+\alpha)$, and discount factor $\bar{\alpha} = \alpha/(1+\alpha)$, where we discount future cost at time $n$ with factor $(1-\bar{\alpha})^n$, $\delta \in \mathcal{D}$. For the associated $\bar{\alpha}$-discount cost value function $\bar{v}^{\bar{\alpha}}(N)$ it holds that $\bar{v}^{\bar{\alpha}}(N) = v^\alpha(N)$. Again, the existence of one policy with finite $\bar{\alpha}$-discounted cost, ensures that $\bar{v}^{\bar{\alpha}}(N)$ is a solution to the $\bar{\alpha}$-discount optimality equation

$$\begin{aligned}u_x &= x_1\bar{c} + (1-\bar{\alpha})\Big(\lambda\big\{\mathbb{1}_{\{x_1>0\}} u_{x-e_1+e_2} + \mathbb{1}_{\{x_1=0\}}(u_{x+e_2} + K^{\text{on}})\big\} \\ &\quad + \{x_2 \wedge N\}\mu \min\{u_{x+e_1-e_2}, u_{x-e_2} + K^{\text{off}}\} + (N-x_2)^+\mu u_x\Big), \qquad (4.6)\end{aligned}$$

and minimising policies are $\bar{\alpha}$-discount optimal in $\Pi$. Furthermore, VI converges to the value function $\bar{v}^{\bar{\alpha}}(N)$, because $\cup_x \mathcal{D}_x$ is finite (cf. [70]), for $N < \infty$. The VI scheme is given by initialising

$$\bar{v}_x^{\bar{\alpha},0}(N) = 0, \quad x \in \mathbf{S}, \qquad (4.7)$$

and then iterating

$$\begin{aligned}&\bar{v}_x^{\bar{\alpha},n+1}(N) \\ &= x_1\bar{c} + (1-\bar{\alpha})\Big(\lambda\big\{\mathbb{1}_{\{x_1>0\}}\bar{v}^{\bar{\alpha},n}_{x-e_1+e_2}(N) + \mathbb{1}_{\{x_1=0\}}(\bar{v}^{\bar{\alpha},n}_{x+e_2}(N) + K^{\text{on}})\big\} \\ &+ \{x_2 \wedge N\}\mu\min\{\bar{v}^{\bar{\alpha},n}_{x+e_1-e_2}(N), \bar{v}^{\bar{\alpha},n}_{x-e_2}(N) + K^{\text{off}}\} + (N-x_2)^+\mu\bar{v}^{\bar{\alpha},n}_x(N)\Big). \\ &\qquad (4.8)\end{aligned}$$

Now, if the functions $\bar{v}^{\bar{\alpha},n}(N)$, $n = 0, 1, \ldots$, are supermodular and superconvex, then the limit $\bar{v}^{\bar{\alpha}}(N) = v^{\alpha}(N)$ is supermodular and superconvex, and thus $v^{\alpha} = \lim_{N\to\infty} v^{\alpha}(N)$ is supermodular and convex. Application of Lemma 4.3.1 then yields the existence of an SC $\alpha$-discount optimal policy. This is summarised in the following theorem.

**Theorem 4.3.2.** *Let $0 < \bar{\alpha} < 1$, $N < \infty$. For the functions $\bar{v}^{\bar{\alpha},n}(N)$ obtained from the value iteration scheme Eqs. (4.7) and (4.8) it holds that*

$$\bar{v}^{\bar{\alpha},n}(N) \in \mathcal{S}uper\mathcal{C} \cap \mathcal{S}uper, n = 0, 1, \ldots. \tag{4.9}$$

*As a consequence, $v^{\alpha} \in \mathcal{S}uper\mathcal{C} \cap \mathcal{S}uper$. Hence, there exists an* SC *$\alpha$-discount optimal policy and the first part of Theorem 4.2.1 holds true.*

*Proof.* Two steps need to be validated. The first is to prove that $v^{\alpha}(N) \to v^{\alpha}$, $N \to \infty$. Essentially this follows from Chapter 3. Theorem 3.5.1 provides conditions under which this continuity result holds and also that $v^{\alpha}$ is finite. These conditions are drift conditions. together with continuity assumptions on the input parameters. Then Theorem 3.6.1 shows that the collection of parametrised processes $\{X(N, \delta)\}$ satisfies the conditions of Theorem 3.5.1. Taking $K^{\text{off}} > 0$ in the present chapter is trivially incorporated.

The second is to prove Eq. (4.9). This is quite a long and tedious proof using event based dynamic programming (cf. [45]). It is therefore postponed to Section 4.5. □

**Average optimality** The case of average optimality of an SC-policy has been open so far. Now, remind that [1] shows average optimality of an SC-policy for the uniformised $N$-perturbations, via a vanishing discount approach based on $\alpha$-discount optimality of an SC-policy for the uniformised $N$-perturbations. They show that the $N$-perturbed server farm model satisfies conditions on the input parameters of the MDP developed in [74] to justify the vanishing discount approach. These conditions are very similar to conditions developed by Borkar in [20, 21], and weakened in [63] (the 'BOR' conditions in this book). Clearly, the vanishing discount then also applies to the continuous time $N$-perturbed MDP.

One approach could be to take the limit $N \to \infty$. Denoting the minimum $N$-perturbed average cost by $\mathsf{g}(N)$, one would in any case need to show that $\mathsf{g}(N) \to \mathsf{g}$, $N \to \infty$. However, the closed class structure under the various policies is rather complicated in the server farm model, and so we have no clue as to what kind of conditions on the input parameters could validate such a continuity result.

The best road therefore seems to be to apply a vanishing discount approach to the continuous time $\alpha$-discount cost value functions. Interestingly, the continuous time versions of the 'BOR' conditions allow to apply the vanishing discount approach to unbounded jump MDPs of Chapter 2. This is expressed in the following theorem. Denote by $\delta^\alpha$ an $\alpha$-discount optimal policy.

**Theorem 4.3.3.** *Any sequence $\{\alpha_n\}_n$, $\alpha_n \downarrow 0$, $n \to \infty$, has a subsequence, again denoted $\{\alpha_n\}_n$, along which the following limits exist: for all $x \in \mathbf{S}$*

$$\begin{aligned} w_x &= \lim_{n\to\infty} (v_x^{\alpha_n} - v_0^{\alpha_n}) \\ g &= \lim_{n\to\infty} \alpha_n v_x^{\alpha_n}. \\ \delta^* &= \lim_{n\to\infty} \delta^{\alpha_n}. \end{aligned}$$

*The tuple $(g, w)$ is a solution to the average optimality equation,*

$$\begin{aligned} g' &= x_1 c + \lambda\Big(\mathbb{1}_{\{x_1>0\}} w'_{x-e_1+e_2} + \mathbb{1}_{\{x_1=0\}}(w'_{x+e_2} + K^{\mathrm{on}})\Big) \\ &\qquad + x_2\mu \min\{w'_{x+e_1-e_2}, w'_{x-e_2} + K^{\mathrm{off}}\} - (\lambda + x_2\mu) w'_x. \quad (4.10) \end{aligned}$$

*Furthermore, $g = \mathsf{g}$, $\delta^*$ is a minimising policy for solution tuple $(\mathsf{g}, w)$ and any minimising policy $\delta$ for solution tuple $(\mathsf{g}, w)$ is average cost optimal in $\Pi$.*

**Corollary 4.3.4.** *For the solution tuple $(\mathsf{g}, w)$ it holds that $w \in \mathcal{S}uper\mathcal{C} \cap \mathcal{S}uper$. Thus, there exists an* SC *average cost optimal policy. This proves the second part of Theorem 4.2.1.*

*Proof of Theorem 4.3.3.* We wish to apply Theorem 2.3.4. On top of the conditions that had to be verified for the assertion of Theorem 4.3.2, the following conditions have to be checked:

1. There exists a policy $\delta^0 \in \mathcal{D}$ and state a $x_0$, such that $m_{xx_0}(\delta^0) < \infty$ and $c_{xx_0}(\delta^0) < \infty$, $x \in \mathbf{S}$, where

$$m_{xx_0}(\delta^0) = \mathsf{E}_x^{\delta^0} \tau_{x_0}, \quad c_{xx_0}(\delta^0) = \mathsf{E}_x^{\delta^0} \int_{t=0}^{\tau_{x_0}} c_{X_t}(D_t) dt,$$

   and

$$\tau_{x_0} = \inf\{t > 0 \,|\, \exists s \in (0, t), \text{ such that } X_s \neq x_0, X_t = x_0\}. \quad (4.11)$$

2. There exists $\epsilon > 0$ such that $A = \{x \in \mathbf{S} \,|\, \min_{\delta_x \in \mathcal{D}_x} c_x(\delta_x) \leq \mathsf{g}(\delta^0) + \epsilon\}$ is a finite set.

3. For all $x \in A \setminus \{0\}$, there exists $\delta^x \in \mathcal{D}$, such that $m_{x_0 x}(\delta^x) < \infty$ and $c_{x_0 x}(\delta^x) < \infty$.

First note that Condition 2 is not true. However, we can use the same trick as in [1] to adapt the cost rate, without influencing the difference in cost of any two policies. Let $\tilde{c}_x(\delta) = c_x(\delta) + x_2$. Denote by $\tilde{\mathsf{g}}(\delta)$ the associated expected average cost, then

$$\tilde{\mathsf{g}}(\pi) = \mathsf{g}(\pi) + \sum_{i=0}^{\infty} i\Big(\frac{\lambda}{i!\mu}\Big)^i e^{-\lambda/\mu} = \mathsf{g}(\pi) + \frac{\lambda}{\mu},$$

since $\{X_{t,2}\}_t$ behaves as an M/M/$\infty$-queue. Thus, enlarging the cost rates in this way has no effect on the optimality characteristics. Then $\tilde{c}_x(\delta_x)$, $x \in \mathbf{S}$, $\delta_x \in \mathcal{D}_x$, satisfies Condition 2.

We will prove Condition 1. Put $x_0 = (0,0) =: 0$. Take $\delta^0_x = \textsc{off}$ if $x \in B$, and $\delta^0_x = \textsc{idle}$ if $x \notin B$, $x_2 > 0$. Then, the set $B \cup \{0\}$ is a closed, irreducible class in the associated Markov process, with finite expected average cost. Thus, $m_{x0}(\delta^0)$, $c_{x0}(\delta^0) < \infty$, for $x \in B$.

Let $x \notin B \cup \{0\}$. Then, since $\{X_{t,2}(\delta^0)\}_t$ behaves as an M/M/$\infty$-queue,

$$m_{x(0,x_1+x_2)}(\delta^0),\ c_{x(0,x_1+x_2)}(\delta^0) < \infty.$$

Thus, also $m_{x,0}(\delta^0) = m_{x,(0,x_1+x_2)}(\delta^0) + m_{(0,x_1+x_2),0}(\delta^0) < \infty$. This applies analogously to $c_{x,0}(\delta^0)$.

We finally have to check Condition 3. If $x \in B$, then we can take $\delta^x = \delta^0$, and a similar reasoning applies. Suppose that $x \notin B$. Then put $\delta^x_y = \textsc{off}$, $y \in B$, $y \neq (0, x_1 + x_2)$, and $\delta^x_y = \textsc{idle}$ for $y$ with $y_1 + y_2 = x_1 + x_2$ and $y_2 > 0$ (and arbitrary in the other states). Again a similar reasoning as in the above applies.

It is easy to verify that the enlarged cost rates do not violate the conditions of Chapter 3 to hold. Thus, for each $\alpha > 0$, $\tilde{v}^\alpha$ is a solution to the $\alpha$-discount optimality equation (4.3), with the right-hand side enlarged by a cost term $x_2$, with the property that any minimising policy is $\alpha$-discount optimal.

Notice that

$$\tilde{v}^\alpha_x - v^\alpha_x = \tilde{m}^\alpha_{x_2}, \quad x \in \mathbf{S},$$

where $\tilde{m}^\alpha$ is the $\alpha$-discount value function of the M/M/$\infty$-queue with cost rate $i$ in state $i$, for $i \in \mathbb{Z}_+$. Checking the conditions of Theorem 3.5.1 (see also Example 3.4.1 'containing' the verification of these conditions), it follows that $\tilde{m}^\alpha$ is finite for $\alpha > 0$. Then $\tilde{v}^\alpha \in \mathcal{S}uper\mathcal{C} \cap \mathcal{S}uper$ for $\alpha > 0$. By Theorem 2.3.4, the assertions in the statement of the present theorem apply to $\tilde{v}^\alpha - \tilde{v}^\alpha_{x_0} e$, with $e$ the vector consisting of ones. Thus any limit point, say $\tilde{w}$, is supermodular and superconvex, and any minimising policy is average cost optimal. Therefore, there exists an SC minimising policy. By the observation made above that the optimality structure is not affected, the original problem has an SC average cost optimal policy.

The assertion of the Theorem is slightly stronger. Now, notice that the M/M/$\infty$-queue is positive recurrent and has finite expected average cost $\lambda/\mu$, for cost rates $i$ in state $i$, $i \in \mathbf{Z}_{\geq 0}$. The conditions of Theorem 2.3.4 therefore hold and so any sequence $\{\alpha_n\}_n$, $\alpha_n \downarrow 0$, $n \to \infty$, contains a subsequence, again denoted by $\{\alpha_n\}_n$, such that $\lim_{n\to\infty}(\tilde{m}^{\alpha_n}_x - \tilde{m}^{\alpha_n}_0) \to m_x$, for some function $m : \mathbf{Z}_{>0} \to \mathbb{R}$, and $\alpha^n \tilde{m}^\alpha_0 \to \lambda/\mu$. Put $v^\alpha_x - v^\alpha_{x_0} = \tilde{v}^\alpha_x - \tilde{v}^\alpha_{x_0} - (\tilde{m}^\alpha_{x_2} - \tilde{m}^\alpha_0)$. The assertion of the Theorem readily follows.

*Proof of Corollary 4.3.4.* Clearly, $w \in \mathcal{S}uper\mathcal{C} \cap \mathcal{S}uper$, as a limit of supermodular, superconvex functions. The same arguments proving Lemma 4.3.1 yield that there exists a minimising policy in Eq. (4.10) with solution tuple tuple $(\mathsf{g}, w)$, that is an SC-policy.

## 4.4 Upper bounds by coupling techniques

In this section we will exploit the special features of the server farm model mentioned directly below Theorem 4.2.2, namely that $\{X_{t,2}\}_t$ behaves as an M/M/$\infty$-queue that is independent of the policy, and that $\Delta(n)$ has only one exit state. These features allow to use coupling techniques. The analysis is first performed for the discounted cost criterion, lateron the limit $\alpha \downarrow 0$ is taken to obtain results for the average cost case. Alternatively, the latter result can be obtained by an analogous derivation as in $\alpha$-discounted cost case.

With this coupling technique we provide upper bounds on the number of states where idling might be optimal. These bounds are expressed in terms of input parameters of the problem. Using that there is an SC $\alpha$-discounted optimal policy by Theorem 4.3.2, it follows that only a finite number of states is left for which the SC-optimal policy still will have to be determined. By virtue of Lemma 4.2.3, it is sufficient to only solve a finite state, finite action MDP. Furthermore, this allows to deduce the existence of a strong Blackwell optimal policy. The original ideas for this section are from Kappetein [42] and Van der Velde [73].

To start with, we will prove Lemma 4.2.3. Denote by $w(n)$ the minimum total expected cost vector for the $n$-restricted MDP.

**Lemma 4.4.1.** *For $x \in \Delta(n) \setminus \{(0,n)\}$ it holds that*

$$v_x^\alpha = v_x^\alpha(n) + \prod_{l=x_2}^{n-1} t_l \cdot v_{(0,n)}^\alpha,$$

*and that*

$$w_x = w_x(n) + \mathsf{E}_{x_2}\tau_n \cdot \mathsf{g},$$

*where* $(\mathsf{g}, w)$ *is a vanishing discount solution of Eq. (4.10) from Theorem 4.3.3.*

*Proof.* We consider the $\alpha$-discounted cost first. A straightforward computation shows for any policy $\pi \in \Pi$ and $x \in \Delta(n) \setminus \{(0,n)\}$ that

$$\begin{aligned} v_x^\alpha(\pi) =& \mathsf{E}_x^\pi \int_{t=0}^{\tau_{(0,n)}} e^{-\alpha t} c_{X_t}(D_t) dt + \mathsf{E}_x^\pi e^{-\alpha \tau_{(0,n)}} v_{(0,n)}^\alpha(\pi) \\ =& v_x^\alpha(n) + \mathsf{E}_x^\pi e^{-\alpha \tau_{(0,n)}} v_{(0,n)}^\alpha(\pi). \end{aligned}$$

Since $\tau_{(0,n)}$ is exclusively determined by the process $\{X_{t,2}\}_t$ if the initial state belongs to $\Delta(n)$, the distribution of $\tau_{(0,n)}$ is equal to the distribution of $\tau_n$ in the M/M/$\infty$-queue, which is independent of the policy. Thus, using [28]

$$\mathsf{E}_x^\pi e^{-\alpha \tau_{(0,n)}} = \mathsf{E}_{x_2} e^{-\alpha \tau_n} = \prod_{l=x_2}^{n-1} t_l,$$

where the empty product $\prod_{l=n}^{n-1} t_l$ is assumed to equal 1.

Next we consider the average cost criterion. Note that $(0,n)$ is reached with probability 1 under any stationary policy on $\mathcal{D}(0,n)$. Consider the limits $w_x = \lim_{m\to\infty}(v_x^{\alpha_m} - v_{(0,n)}^{\alpha_m})$ and $\mathsf{g} = \lim_{m\to\infty} \alpha_m v_{(0,n)}^{\alpha_m}$, $\alpha_m \downarrow 0$, $x \in \mathbf{S}$, along some subsequence $\{\alpha_m\}_m$. By virtue of Theorem 4.3.3 this exists, and yields a solution to the average optimality equation (4.10).

In terms of the $n$-restriction, we get

$$v_x^\alpha(n) = v_x^\alpha - v_{(0,n)}^\alpha + \frac{1 - \mathsf{E}_{x_2} e^{-\alpha\tau_n}}{\alpha} \alpha v_{(0,n)}^\alpha \to w_x + \mathsf{E}_{x_2}\tau_n \cdot \mathsf{g},$$

where $\mathsf{E}_n \tau_n = 0$. Denote $w_x(n) = w_x + \mathsf{E}_{x_2}\tau_n \cdot \mathsf{g}$, $x \in \Delta(n)$. Then, plugging this into Eq. (4.10) for $x \in \Delta(n)$ yields the equation

$$\begin{aligned} 0 =& x_1 c + \lambda \Big\{ \mathbb{1}_{\{x_1>0\}} w_{x-e_1+e_2}(n) + \mathbb{1}_{\{x_1=0\}} (w_{x+e_2}(n) + K^{\text{on}}) \Big\} \\ &+ x_2\mu \cdot \min\{w_{x+e_1-e_2}(n), w_{x-e_2}(n) + K^{\text{off}}\} \\ &- (\lambda + x_2\mu) w_x(n) \qquad\qquad x \in \Delta(n) \setminus \{(0,n)\}. \end{aligned} \tag{4.12}$$

Thus, computing the average cost value function on $\Delta(n)$ is equivalent to solving a total cost problem till absorption into state $(0,n)$, given the optimal average cost $\mathsf{g}$ is known. $\square$

*Proof of Lemma 4.2.3.* Note that $v_x^\alpha(n) = v_x^\alpha - \prod_{l=x_2}^{n-1} t_l \cdot v_{(0,n)}^\alpha$, $x \in \Delta(n)$, is the unique solution to the $\alpha$-discount optimality equation for the $n$-restriction on $\Delta(n)$. Thus, any minimising policy is optimal for both the unrestricted and the restricted problems on $\Delta(n)$. Any $\alpha$-discount optimal policy must satisfy the $\alpha$-discount optimality equation, for both problems. For the $n$-restriction this is immediate, by uniqueness of the solution to the $\alpha$-discount optimality equation. For the unrestricted problem this follows from Theorem 3.4.1.

Furthermore, a stationary optimal policy for the $n$-restriction chooses minimising decisions in Eq. (4.12), hence in Eq. (4.10), for $x \in \Delta(n) \setminus \{(0,n)\}$. $\square$

**$\alpha$-Discounted cost criterion** In view of Lemma 4.2.3, we can study the characteristics of optimal policies on the triangle $\Delta(n)$ for the $n$-restricted MDP. To recall, $v^\alpha(n) : \Delta(n) \to \mathbb{R}$ denotes the $\alpha$-discount value function. It is the unique solution to

$$\alpha u_x = \begin{cases} cx_1 + \lambda\Big\{\mathbb{1}_{\{x_1>0\}} u_{x-e_1+e_2} + \mathbb{1}_{\{x_1=0\}}(u_{x+e_2} + K^{\mathrm{on}})\Big\} & -(\lambda + x_2\mu)u_x \\ \quad +x_2\mu \cdot \min\{K^{\mathrm{off}} + u_{x-e_2}, u_{x-e_2+e_1}\} & x \in \Delta(n) \setminus \{(0,n)\}; \\ 0 & x = (0,n). \end{cases} \tag{4.13}$$

For studying the characteristics of optimal policies, we will view the $n$-restricted MDP as a (continuous time) semi-Markov decision process (SMDP), in particular an exponential semi-Markov decision process (ESMDP), since the time between jumps has an exponential distribution. The advantage is that we may consider non-stationary deterministic policies, that only at the time of a jump prescribe a decision, based on the history of past states and decisions at jump times. By right-continuity of trajectories, the decision is taken in the just reached state. Following [57, Chapter 7], we conclude that there exists a stationary, deterministic $\alpha$-discounted optimal policy, since Condition 1 from that book is trivially satisfied in our case. We will use this in our coupling arguments.

Denote the class of non-stationary, deterministic policies for the ESMDP by $\mathcal{C}$. Then again, the $\alpha$-discount cost value function associated with policy $R \in \mathcal{C}$ is denoted by $v^\alpha(n, R)$. Let $x \in \Delta(n)$. Let $R, R' \in \mathcal{C}$. Then one can construct a probability space, such that for the associated right-continuous non-stationary processes $X(R)$ and $X(R')$ it holds that

$$X_{t,2} := X_{t,2}(R) = X_{t,2}(R'), \quad t > 0,$$

and $X_0(R) = X_0(R') = x$. Then the two processes jump at the same time, but may take different decisions. For notational convenience, denote the corresponding probability and expectation operators by $\mathsf{P}^*$ and $\mathsf{E}^*$ respectively.

Denote the successive jump times by $J_0 = 0, J_1, J_2, \ldots$. Thus at time $J_n$ the states of the processes (after the jump, by right-continuity) are given by $X_{J_n}(R)$ and $X_{J_n}(R')$ respectively. The corresponding decisions are denoted $D_{J_n}(R)$ and $D_{J_n}(R')$ respectively. This construction will be assumed to apply during the whole course of this section.

Let $X_{J_0}(R), X_{J_0}(R') \in \Delta(n) \setminus \{(0,n)\}$. Suppose that there exists a random time $\sigma_x$, such that

- $\sigma_x \in \{J_1, \ldots\}$;
- $X_{\sigma_x}(R) = X_{\sigma_x}(R')$;
- for $J_m \geq \sigma_x$, $X_{J_m}(R) = X_{J_m}(R')$ and $D_{J_m}(R) = D_{J_m}(R')$.

Then $\sigma_x$ is called a *coupling time* for $X(R)$ and $X(R')$. Notice that $\sigma_x$ is almost surely finite, since the processes couple latest upon reaching state $(0,n)$. This happens in a.s. finite time by positive recurrence of the M/M/$\infty$-queue.

If $\sigma_x$ is a coupling time for $X(R)$ and $X(R')$ then

$$v_x^\alpha(n,R) \leq v^\alpha(n,R') \quad \Longleftrightarrow \quad v_{x,\sigma_x}^\alpha(n,R) \leq v_{x,\sigma_x}^\alpha(n,R'), \tag{4.14}$$

where

$$v_{x,\sigma_x}^\alpha(n,R) = \mathsf{E}_x^R \int_0^{\sigma_x} e^{-\alpha t} c_{X_t}(D_t)dt,$$

and analogously for $v_{x,\sigma_x}^\alpha(n,R')$.

The following simple lemma validates the coupling arguments that we will use.

**Lemma 4.4.2.** *Let $x \in \mathbf{S}^* \cap \Delta(n) \setminus \{(0,n)\}$, and $d_1, d_2 \in \mathcal{D}_x$ with $d_1 \neq d_2$. Suppose that two policies $R_1, R_2$ are such that:*

1. *$X_0(R_1) = X_0(R_2) = x$;*
2. *$R_1$ is a stationary deterministic policy on $\Delta(n)$ that takes decision $d_1$ in $x$ and optimal decisions in the states of $\Delta(n) \setminus \{x\}$;*

3. $D_{J_0}(R_2) = d_2$;

4. *there exist deterministic policies $R_{2a}$ and $R_{2d}$ depending on whether at time $J_1$ an arrival or a departure respectively took place, describing the decisions of $R_2$ after $J_1$;*

5. *there exists a coupling time $\sigma_x$;*

6. $v^\alpha_{x,\sigma_x}(n, R_1) \geq v^\alpha_{x,\sigma_x}(n, R_2)$.

*Then, $d_2$ is an optimal in state $x$, that is, it is a minimising decision in Eq. (4.13). It is the unique minimising decision in Eq. (4.13) if there is strict inequality in condition (6).*

*Proof.* If $d_2$ is not an optimal decision, then $d_1$ must be optimal. Hence $R_1$ is optimal, and is equal to say the deterministic, stationary policy $\delta$ say (on $\Delta(n)$). Let $R$ be the policy with $D_{J_0}(R) = d_2$, and $D_{J_m}(R) = \delta_{X_{J_m}(R)}$, $m \geq 1$. That is, after the first decision, $R$ equals the $\alpha$-discounted optimal policy $R_1$. Then,

$$v^\alpha_x(n, R_1) < v^\alpha_x(n, R).$$

This is immediately deducible from the $\alpha$-discounted optimality equation. Thus,

$$\begin{aligned} v^\alpha_x(n, R_1) <& v^\alpha_x(n, R) \\ =& \frac{c_x(d_2)}{\lambda + x\mu + \alpha} + \frac{\lambda}{\lambda + x\mu + \alpha} v^\alpha_{x'}(n, R_1) + \frac{x\mu}{\lambda + x\mu + \alpha} v^\alpha_{x''}(n, R_1) \\ \leq& \frac{c_x(d_2)}{\lambda + x\mu + \alpha} + \frac{\lambda}{\lambda + x\mu + \alpha} v^\alpha_{x'}(n, R_{2a}) + \frac{x\mu}{\lambda + x\mu + \alpha} v^\alpha_{x''}(n, R_{2d}) \\ =& v^\alpha_x(n, R_2), \end{aligned}$$

where $x'$ and $x''$ are the resulting states after an arrival and a departure respecively when decision $d_2$ is implemented. This contradicts condition (6) in the assertion of the lemma by virtue of Eq. (4.14). □

For convenience, introduce $T_i \stackrel{d}{=} \exp(\lambda + i\mu)$, a random variable that represents the sojourn time in state $x$, when $x_2 = i$.

**Lemma 4.4.3.** *1. If $c \geq (\lambda + \alpha)K^{\text{off}} + \lambda K^{\text{on}}$ then $\pi(0, 1) =$ OFF is $\alpha$-discount optimal;*

2. *if $c \leq (\lambda + \alpha)K^{\text{off}} + \lambda K^{\text{on}}$ then $\pi(0, 1) =$ IDLE is $\alpha$-discount optimal.*

*Proof.* We will prove the second statement. It is sufficient to consider the 2-restricted processes on $\Delta(2)$. Start in state $x = (0,1)$. Put $d_1 = \text{OFF}$, $d_2 = \text{IDLE}$. Let $\delta$ be a deterministic, stationary $\alpha$-discount optimal policy on $\Delta(2)$.

$R_1$ is the stationary, deterministic policy $\delta^1$, with $\delta^1_x = d_1 = \text{OFF}$, and $\delta^1_y = \delta_y$, $y \neq x$. $R_2$ is the stationary, deterministic policy $\delta^2$ with $\delta^2_x = d_2 = \text{IDLE}$, and $\delta^2_y = \delta_y$, $y \neq x$.

Notice that the time to couple $\sigma_x$ is equal to $J_1$ with probability $\lambda/(\mu+\lambda)$, and $J_2$ with probability $\mu/(\mu+\lambda)$. We get

$$\begin{aligned}
v^\alpha_{x,\sigma_x}(2,R_1) - v^\alpha_{x,\sigma_x}(2,R_2) =& \mathsf{E}^*_x\Big\{\int_0^{J_1} e^{-\alpha t}\mu K^{\text{off}} dt\Big\} \\
&+ \frac{\mu}{\lambda+\mu}\mathsf{E}^*_x\Big\{e^{-\alpha J_1}\int_0^{J_2-J_1} e^{-\alpha t}(\lambda K^{\text{on}} - c)dt \,\Big|\, T>1\Big\} \\
=& \frac{\mu}{\alpha+\mu+\lambda}\Big\{K^{\text{off}} + \mathsf{E}\int_{t=0}^{T_0} e^{-\alpha t}(\lambda K^{\text{on}} - c)dt\Big\} \\
=& \frac{\mu}{\alpha+\mu+\lambda}\,\frac{(\lambda+\alpha)K^{\text{off}} + \lambda K^{\text{on}} - c}{\alpha+\lambda} \geq 0.
\end{aligned}$$

By virtue of Lemma 4.4.2, the IDLE decision is a minimising decision in Eq. (4.13) for $x = (0,1)$, hence it is in Eq. (4.3). □

*Proof of Theorem 4.2.2.* First we prove assertion (1,i). To this end, assume $c > (\lambda+\alpha)K^{\text{off}} + \lambda K^{\text{on}}$. Then, Lemma 4.4.3 (1) implies that OFF is optimal in $(0,1)$. By virtue of Theorem 4.2.1 (i), turning off is optimal in all states.

Next we will prove assertion (1,ii). Assume $c \leq (\lambda+\alpha)K^{\text{off}} + \lambda K^{\text{on}}$. We first show the validity of (1,ii,a) that the decision IDLE is a minimising decision in Eq. (4.3) for $x \in B$, with $B$ the $x_2$-axis minus the origin.

The proof is by induction. Lemma 4.4.3 (2) implies that IDLE is optimal in state $(0,1)$. Now assume that IDLE is optimal in $\{(0,1),\ldots,(0,n)\}$.

We will compare two policies $R_1$ and $R_2$, for initial state $x = (0,n+1)$. It is sufficient to study the $(n+2)$-restricted MDP on $\Delta(n+2)$. Let $\delta$ be a deterministic $\alpha$-discount optimal policy on $\Delta(n+2)$, with $\delta_{(0,i)} = \text{IDLE}$, $i = 1,\ldots,n$. Put $d_1 = \text{OFF}$, $d_2 = \text{IDLE}$. Let $R_1$ be the deterministic, stationary policy $\delta^1$ that selects decision $d_1$ in state $x$ and $\delta_y$ in state $y \neq x$, $x \in \Delta(n+2)$. The policy $R_2$ is specified by $d_2$ at time $J_0$, and policies $R_{2a}$, and $R_{2d}$ depending on whether there was an arrival at time $J_1$ or a departure. $R_{2a}$ will be equal to $\delta^1$. $R_{2d}$ is given by

$$D_{J_0}(R_{2d}) = \begin{cases} 0 & X_{J_0}(R_{2d}) = (0, n+2) \\ \textsc{off} & X_{J_0}(R_{2d}) = (1, n) \\ \text{arbitrary} & X_{J_0}(R_{2d}) = y, \qquad\qquad y \neq (0, n+2), (1, n), \end{cases}$$

and for $m \geq 1$, $D_{J_m}(R_{2d}) = \delta^1_{X_{J_m}(R_{2d})}$. Then the coupling time $\sigma_x$ is equal $J_1$ to with probability $\lambda/(\lambda + (n+1)\mu)$ and occurs at state $(0, n+2)$. It is $J_2$ with probability $(n+1)\mu/(\lambda+(n+1)\mu)$, and coupling occurs with subsequent probability $p := \lambda/(\lambda + n\mu)$ at state $(0, n+1)$ due to an arrival, and with probability $1 - p = n\mu/(\lambda + n\mu)$ at state $(1, n-1)$ due to a departure. We get

$$\begin{aligned} & v^{\alpha}_{x,\sigma_x}(n+2, R_1) - v^{\alpha}_{x,\sigma_x}(n+2, R_2) \\ =& \mathsf{E}^*_x\Big\{\int_0^{J_1} e^{-\alpha t}(n+1)\mu K^{\text{off}} dt\Big\} + \\ & \quad \frac{(n+1)\mu}{(n+1)\mu + \lambda}\mathsf{E}^*_x\Big\{e^{-\alpha J_1}\int_0^{J_2 - J_1} e^{-\alpha t}(\lambda K^{\text{on}} - c - n\mu K^{\text{off}})dt \,\Big|\, T > 1\Big\} \\ =& \frac{(n+1)\mu}{\alpha + (n+1)\mu + \lambda}\frac{(\lambda+\alpha)K^{\text{off}} + \lambda K^{\text{on}} - c}{\alpha + n\mu + \lambda} > 0. \end{aligned}$$

Thus, decision $d_2 = \textsc{idle}$ is optimal in state $x = (0, n+1)$. By induction $\textsc{idle}$ is optimal in the states of $B$.

In order to prove assertion (1,ii,b), we will derive the following preliminary claim.

**Claim.** *If $c > \alpha K^{\text{off}}$ then for all $n \geq \{2 \wedge n_1(\alpha)\}$ it holds that* $\textsc{off}$ *is (strictly) optimal in state $(n-1, 1)$.*

*Proof of the Claim.* Supposing that $c > \alpha K^{\text{off}}$, let $n \geq 2 \wedge n_1(\alpha)$. It is sufficient to consider the $n$-restricted MDP on $\Delta(n)$. Let $x = (n-1, 1)$, and let $d_1 = \textsc{idle}$. Let $\delta$ be an SC $\alpha$-discount optimal policy on $\Delta(n)$. If $\delta_x = \textsc{off}$ then there is nothing to prove. Hence, assume that $\delta_x = \textsc{idle}$. Then $\delta_y = \textsc{idle}$ for $y \in \mathbf{S}^* \cap \Delta(n) \setminus \{(0, n)\}$. The policy $R_1$ is then equal to $\delta$, thus, $R_1$ idles in every state of $\Delta(n)$.

Put $d_2 = \textsc{off}$. Subsequently, we put $R_{2a}$ and $R_{2d}$ equal to $\delta$. Then, either $X(R_1)$ and $X(R_2)$ couple at state $(n-2, 2)$ at time $J_1$ due to an arrival, or they couple at state $(0, n)$, and till that moment (on the event $T > 1$) the process $X(R_2)$ moves on the set of states $\{y \,|\, y_1 + y_2 = n - 1\}$. Note that $T$ is finite with probability 1, since state $(0, n)$ is reached with probability 1 and in finite expected time. That is, $\sigma_x = J_T$ is finite with probability 1.

The cost difference between the two processes consists of a switch-off cost in process $X(R_2)$ during the interval $[0, J_1)$. In case of a service completion at time $J_1$, there is an additional switch-on cost in process $X(R_2)$ at time intervals during $[J_1, J_T)$ that no idling servers are present, as well as idling cost in process $X(R_1)$ during all of $[J_1, J_T)$. Let

$$m_0 = \min\{l \mid X_{J_l}(R_2) = (0, n-1)\} = \min\{l \mid X_{J_l,2} = 0\}.$$

We thus get,

$$\begin{aligned}
&v^{\alpha}_{x,\sigma_x}(n, R_1) - v^{\alpha}_{x,\sigma_x}(n, R_2)\\
=& -\mu K^{\mathrm{off}} \mathsf{E}^*_x \int_{t=0}^{J_1} e^{-\alpha t} dt\\
&+ \frac{\mu}{\lambda+\mu} \mathsf{E}^*_x \Big\{ c \int_{J_1}^{J_T} e^{-\alpha t}\, dt - K^{\mathrm{on}} \int_{J_{m_0}}^{J_T} e^{-\alpha t} \lambda \cdot \mathbb{1}_{\{(0,n-1)\}}(X_t(R_2))\, dt \Big| T > 1 \Big\}\\
=& \frac{\mu}{\alpha+\mu+\lambda} \Big[ -K^{\mathrm{off}} + c \cdot \mathsf{E}^*_x \Big\{ \int_0^{J_T - J_1} e^{-\alpha t}\, dt \Big| T > 1 \Big\}\\
&- K^{\mathrm{on}} \cdot \mathsf{E}^*_x \Big\{ e^{-\alpha(J_{m_0} - J_1)} \int_0^{J_T - J_{m_0}} e^{-\alpha t} \lambda \mathbb{1}_{\{(0,n-1)\}}(X_t(R_2))\, dt \Big| T > 1 \Big\} \Big].
\end{aligned} \tag{4.15}$$

We will first consider the coefficient of $K^{\mathrm{on}}$ in Eq. (4.15). Let $J_{m_0} < J_{m_1} < \cdots < J_T$ be the successive instants that $X(R_2)$ hits state $(0, n-1)$. To this end, we bound the integral in expectation after further conditioning on the state $X_{J_{m_0}}(R_2) = (0, n-1)$.

$$\begin{aligned}
&\mathsf{E}^*_x \Big\{ \int_0^{J_T - J_{m_0}} e^{-\alpha t} \lambda \cdot \mathbb{1}_{\{(0,n-1)\}}(X_t(R_2))\, dt \,\Big|\, T > 1, X_{J_{m_0}}(R_2) = (0, n-1) \Big\}\\
&\le \mathsf{E}^*_x \Big\{ \frac{\lambda}{\alpha + (n-1)\mu + \lambda}\\
&+ \sum_{l=1}^{\infty} \Big( \frac{(n-1)\mu}{\alpha + (n-1)\mu + \lambda} \Big)^l \frac{\lambda}{\alpha + (n-1)\mu + \lambda} \Big| T > 1, X_{J_{m_0}}(R_2) = (0, n-1) \Big\}\\
&= \frac{\lambda}{\alpha + \lambda}.
\end{aligned} \tag{4.16}$$

The derivation cancels the part of the paths of $X(R_2)$, during time-intervals $[J_{m_0+1}, \ldots, J_{m_1})$, $[J_{m_1+1}, J_{m_2})$, etc. where it moves on $\Delta(n-1) \backslash \{(0, n-1)\}$. Here we use that from state $(1, n-2)$ there is a return to $(0, n-1)$ with probability 1. This enlarges the contribution of the discount factor. It yields the exact expression in case $n = 2$.

Since there must be at least $n-1$ arrivals till state $(0, n-1)$ is hit, $J_{m_0} - J_1 \geq \theta_{n-1}$, with $\theta_{n-1} \stackrel{d}{=} \mathrm{E}(\lambda, n-1)$, i.e. an Erlang distribution with $n-1$ exponentially distributed phases with parameter $\lambda$. Notice that $\mathsf{E} e^{-\alpha\theta_{n-1}} = \big(\lambda/(\alpha+\lambda)\big)^{n-1}$. Thus, Eq. (4.16) implies that

$$\mathsf{E}_x^*\Big\{e^{-\alpha(J_{m_0}-J_1)}\int_0^{J_T-J_{m_0}} e^{-\alpha t}\lambda\mathbb{1}_{\{(0,n-1)\}}(X_t(R_2))\,dt\Big|T>1\Big\} \leq \Big(\frac{\lambda}{\alpha+\lambda}\Big)^n. \tag{4.17}$$

Similarly, in total at least $n$ arrivals must have occurred between times $J_1$ and $J_T$. Thus, for the coefficient of $c$ in Eq. (4.15) it holds, that

$$\begin{aligned}\mathsf{E}_x^*\Big\{\int_0^{J_T-J_1} e^{-\alpha t}\,dt\,\Big|T>1\Big\}\\ &=\mathsf{E}_x^*\Big\{\frac{1-e^{-\alpha(J_T-J_1)}}{\alpha}\,dt\,\Big|T>1\Big\}\\ &\geq\frac{1}{\alpha}\Big\{1-\Big(\frac{\lambda}{\alpha+\lambda}\Big)^n\Big\}.\end{aligned} \tag{4.18}$$

Finally, inserting Eqs. (4.17) and (4.18) into Eq. (4.15) yields

$$v_{x,\sigma_x}^\alpha(n,R_1) - v_{x,\sigma_x}^\alpha(n,R_2) \geq \frac{\mu}{\alpha+\mu+\lambda}\Big[\frac{c-\alpha K^{\mathrm{off}}}{\alpha} - \frac{c+\alpha K^{\mathrm{on}}}{\alpha}\Big(\frac{\lambda}{\alpha+\lambda}\Big)^n\Big]. \tag{4.19}$$

Finally, we have to choose $n$ large enough, so that the right-hand side of Eq. (4.19) is non-negative. In other words, $n$ has to be large enough, so that

$$\frac{c+\alpha K^{\mathrm{on}}}{c-\alpha K^{\mathrm{off}}} \leq \Big(1+\frac{\alpha}{\lambda}\Big)^n.$$

Since

$$\Big(1+\frac{\alpha}{\lambda}\Big)^n > 1+\frac{n\alpha}{\lambda}, \quad n\geq 2,$$

it is sufficient to require that

$$\frac{c+\alpha K^{\mathrm{on}}}{c-\alpha K^{\mathrm{off}}} \leq 1+\frac{n\alpha}{\lambda}.$$

In other words, it is sufficient that $n \geq \{2\wedge n_1(\alpha)\}$. This completes the proof that it is optimal to turn off in $(n-1,1)$ for $n \geq \{2 \wedge n_1(\alpha)\}$.

The Claim allows to prove assertion (ii,b). Thus, assume $\alpha K^{\text{off}} < c$. Let $n \geq n_0(\alpha)$. We consider the $n$-restricted MDP on $\Delta(n)$. Let $x = (1, n-1)$. Let $\delta$ be an SC $\alpha$-discount optimal policy on $\Delta(n)$.

By virtue of the Claim, there exists a state $\bar{x}$, $\bar{x}_1 + \bar{x}_2 = n$, where $\delta_{\bar{x}} = \text{OFF}$, since $n \geq \{2 \wedge n_1(\alpha)\}$. If $\bar{x} = (1, n-1)$, there is nothing to prove. So, assume that $\bar{x}_1 > 1$, but is minimally chosen, in the sense that $\delta_y = \text{IDLE}$, for $x_1 = 1 \leq y_1 < \bar{x}_1$, and $y_1 + y_2 = n$. Note that $\bar{x}_1 \leq n_1(\alpha) - 1$, and so $\bar{x}_2 \geq n - n_1(\alpha) + 1 \geq n - n_1(\alpha)$.

Choose $d_1 = \text{IDLE}$. Let $R_1$ be the stationary, deterministic policy $\delta^1$ (on $\Delta(n)$), with $\delta^1_x = d_1 = \text{IDLE}$ and $\delta^1_y = \delta_y$, $y \in \Delta(n) \setminus \{x\}$. The policy $R_2$ is defined as follows. First, $d_2 = \text{OFF}$. Define for the M/M/$\infty$ queue that models $X_{t,2}(R_i) = X_{t,2}$, $i = 1, 2$, $t \geq 0$

$$T = \min\Big\{m \geq 1 \,|\, X_{J_m,2} \in \{\bar{x}_2 - 1, n\}\Big\}.$$

Choose $R_{2a}$ equal to $\delta^1$. The policy $R_{2d}$ is chosen as follows: $D_{J_m}(R_{2d}) = \text{IDLE}$, $0 \leq m < T$. For $m \geq T$, put $D_{J_m}(R_{2d}) = \delta^1_{X_{J_m}(R_2)}$. Thus $X(R_1)$ and $X(R_2)$ will couple either in state $\bar{x} - e_2$, or in state $(0, n)$, at time $\sigma_x = J_T$. Note that $X_t(R_1) = X_t(R_2) + e_1$, $J_1 \leq t < J_T$. We get,

$$\begin{aligned}
&v^\alpha_{x,\sigma_x}(n, R_1) - v^\alpha_{x,\sigma_x}(n, R_2)\\
&\quad= \frac{(n-1)\mu}{\lambda + (n-1)\mu} \mathsf{E}^*_x \Big\{ c \int_{J_1}^{J_T} e^{-\alpha t} dt \\
&\qquad + K^{\text{off}} \int_{J_1}^{J_T} e^{-\alpha t} \bar{x}_2\, \mu \cdot \mathbb{1}_{\{\bar{x}\}}(X_t(R_1)) dt \,\Big|\, T > 1 \Big\}\\
&\qquad - \mathsf{E}^*_x\, K^{\text{off}} \int_{t=0}^{J_1} e^{-\alpha t} (n-1)\mu dt\\
&\qquad - \frac{(n-1)\mu}{\lambda + (n-1)\mu} \mathsf{E}^*_x \Big\{ K^{\text{on}} \int_0^{J_T - J_1} e^{-\alpha t} \lambda \cdot \mathbb{1}_{\{(0,n-1)\}}(X_t(R_2)) dt \,\Big|\, T > 1 \Big\}\\
&\quad= \frac{(n-1)\mu}{\alpha + (n-1)\mu + \lambda} \Big[ \mathsf{E}^*_x \Big\{ c \int_0^{J_T - J_1} e^{-\alpha t} dt \\
&\qquad + K^{\text{off}} \int_0^{J_T - J_1} e^{-\alpha t} \bar{x}_2\, \mu \cdot \mathbb{1}_{\{\bar{x}\}}(X_t(R_2)) dt \,\Big|\, T > 1 \Big\}\\
&\qquad - K^{\text{off}} - K^{\text{on}} \mathsf{E}^*_x \Big\{ \int_0^{J_T - J_1} e^{-\alpha t} \lambda \cdot \mathbb{1}_{\{(0,n-1)\}}(X_t(R_2)) dt \,|\, T > 1 \Big\} \Big]. \qquad (4.20)
\end{aligned}$$

We only have to consider the expression between the square brackets. It is convenient to introduce some further notation and derive some preparatory bounds. Let $m_0 = \min\{l \,|\, X_{J_l,2} = \bar{x}_2\}$, which is a random variable with potential value $\infty$, because the process may be absorbed in $(0,n)$ before reaching $\bar{x}_2$. The event $\{J_{m_0} < J_T\}$ has the natural interpretation in the M/M/$\infty$ queue as the event that $\bar{x}_2$ is reached before $n$. Define

$$\gamma(\alpha) = \mathsf{E}_{\bar{x}_2+1}\{e^{-\alpha J_{m_0}} \mathbb{1}_{\{J_{m_0}<J_T\}}\}.$$

Then,

$$\gamma(\alpha) \geq \frac{(\bar{x}_2+1)\mu}{\alpha + (\bar{x}_2+1)\mu + \lambda} \geq 1 - \frac{\alpha+\lambda}{\alpha + \bar{x}_2\mu + \lambda}. \tag{4.21}$$

Next, we need

$$\begin{aligned} \Gamma(\alpha) &= \mathsf{E}^*[x]\{e^{-\alpha(J_{m_0}-J_1)} \,|\, T>1, J_{m_0} < J_T\}, \\ p &= \mathsf{P}^*_x\{J_{m_0} < J_T \,|\, T>1\} = \mathsf{P}_{n-2}\{J_{m_0} < J_T\}, \end{aligned}$$

since $\{T>1\} = \{X_{J_1,2} = x_2 - 1 = n-2\}$. The latter probability can be lower bounded in the following way. First observe that $p$ is at least as large as the probability in the M/M/$\infty$ queue to reach state 0 before reaching $n$, starting at state $n-2$. Thus $1-p$ is smaller than or equal to the probability of reaching $n$ before 0, starting at $n-2$. Formally, write for $0<l<n$

$$a(l) = \mathsf{P}_l\{X_{t,2} > 0,\, t \in [0, \tau_{l+1}]\}. \tag{4.22}$$

Recall that $\tau_{l+1} = \inf\{m \geq 1 \,|\, X_{J_m,2} = l+1\}$. Then $1-p \leq a(n-2)a(n-1)$. Lemma 4.6.1 shows that $a(l) \leq k(\rho)/l$, for $k(\rho) = (1+\rho)^2 e^{\rho}/\rho$.

$$1-p \leq a(n-2)a(n-1) \leq k^2(\rho)\frac{1}{(n-2)(n-1)}. \tag{4.23}$$

The switch-off cost associated with $X(R_1)$ can now be lower bounded as follows.

$$\begin{aligned} &K^{\text{off}}\,\mathsf{E}^*_x\Big\{\int_0^{J_T-J_1} e^{-\alpha t}\bar{x}_2\,\mu \cdot \mathbb{1}_{\{\bar{x}\}}(X_t(R_2))dt \,\Big|\, T>1\Big\} \\ &\geq K^{\text{off}}\,\mathsf{E}^*_x\Big\{\int_{J_{m_0}-J_1}^{J_T-J_1} e^{-\alpha t}\bar{x}_2\,\mu \cdot \mathbb{1}_{\{\bar{x}\}}(X_t(R_2))\,\mathbb{1}_{\{J_{m_0}<J_T\}}dt \,\Big|\, T>1\Big\} \\ &\geq K^{\text{off}}\,\mathsf{E}^*_x\{e^{-\alpha(J_{m_0}-J_1)}\mathbb{1}_{\{J_{m_0}<J_T\}} \,|\, T>1\} \\ &\qquad \cdot \frac{\bar{x}_2\mu}{\alpha+\bar{x}_2\mu+\lambda}\sum_{l\geq 0}\Big(\frac{\lambda}{\alpha+\bar{x}_2\mu+\lambda}\gamma(\alpha)\Big)^l \end{aligned}$$

$$
\begin{aligned}
\geq & K^{\rm off} \cdot \Gamma(\alpha) p \Big\{ 1 - \frac{\alpha}{\alpha + \bar{x}_2\mu + \frac{\lambda(\lambda+\alpha)}{\alpha+\bar{x}_2\mu+\lambda}} \\
& \qquad - \frac{\lambda(\lambda+\alpha)}{(\alpha+\bar{x}_2\mu+\lambda)(\alpha+\bar{x}_2\mu+\lambda\frac{\lambda+\alpha}{\alpha+\bar{x}_2\mu+\lambda})} \Big\},
\end{aligned} \tag{4.24}
$$

where we use the Markov property repeatedly for the second inequality and Eq. (4.21) for the third. For the idle cost term we get

$$
\begin{aligned}
& c \cdot \mathsf{E}_x^* \Big\{ \int_0^{J_T - J_1} e^{-\alpha t}\, dt \,|\, T > 1 \Big\} \\
\geq\ & c \cdot \mathsf{E}_x^* \Big\{ \int_0^{\min\{J_{m_0}, J_T\} - J_1} e^{-\alpha t} dt \,|\, T > 1 \Big\} \\
& \qquad + c \cdot \mathsf{E}_x^* \Big\{ \mathbb{1}_{\{J_{m_0} < J_T\}} \int_{J_{m_0} - J_1}^{J_T - J_1} e^{-\alpha t}\, dt \,|\, T > 1 \Big\} \\
\geq\ & c \cdot \frac{1 - \mathsf{E}_{n-2} e^{-\alpha \min\{J_{m_0}, J_T\}}}{\alpha} + c \cdot \frac{\Gamma(\alpha) p}{\alpha + \bar{x}_2\mu + \frac{\lambda(\lambda+\alpha)}{\alpha+\bar{x}_2\mu+\lambda}},
\end{aligned} \tag{4.25}
$$

by splitting the integral in the part till time $\min\{J_{m_0}, J_T\}$, and the restriction of the remaining part of the integral to periods that a switch off cost is incurred (and thus on the event $\{J_{m_0} < J_T\}$), along trajectories that only jump between the states $\bar{x}$ and $\bar{x} - e_1 + e_2$ till absorption in $\bar{x} - e_2$.

Next, using Eq. (4.23)

$$
\begin{aligned}
\mathsf{E}_{n-2} e^{-\alpha \min\{J_{m_0}, J_T\}} &\leq p\Gamma(\alpha) + \big(1 - p\big) \frac{(n-2)\mu + \lambda}{\alpha + (n-2)\mu + \lambda} \\
&\leq p\Gamma(\alpha) + k^2(\rho) \frac{1}{(n-2)(n-1)}.
\end{aligned} \tag{4.26}
$$

Adding up the idle cost terms and the switch off costs within the square brackets in Eq. (4.20) and using the bounds in Eqs. (4.24), (4.25) and (4.26), yields

$$
\begin{aligned}
& \mathsf{E}_x^* \Big\{ c \int_0^{J_T - J_1} e^{-\alpha t} dt \\
& \qquad + K^{\rm off} \int_0^{J_T - J_1} e^{-\alpha t} \bar{x}_2\, \mu \cdot \mathbb{1}_{\{\bar{x}\}}(X_t(R_2)) dt\, \mathbb{1}_{\{T < \tau\}} \Big| T > 1 \Big\} - K^{\rm off} \\
\geq & \Gamma(\alpha) p \frac{c - \alpha K^{\rm off} - \frac{\lambda(\lambda+\alpha)}{\alpha+\bar{x}_2\mu+\lambda} K^{\rm off}}{\alpha + \bar{x}_2\mu + \lambda\frac{\lambda+\alpha}{\alpha+\bar{x}_2\mu+\lambda}} + \frac{1 - \mathsf{E}_{n-2} e^{-\alpha(\min\{J_{m_0}, J_T\} - J_1)}}{\alpha} (c - \alpha K^{\rm off})
\end{aligned}
$$

$$
\begin{aligned}
&+\Big(\Gamma(\alpha)p - \mathsf{E}_{n-2} e^{-\alpha(\min\{J_{m_0},J_T\}-J_1)}\Big)K^{\mathrm{off}} \\
\geq& \Gamma(\alpha)p \frac{c-\alpha K^{\mathrm{off}} - \frac{\lambda(\lambda+\alpha)}{\alpha+\bar{x}_2\mu+\lambda}K^{\mathrm{off}}}{\alpha+\bar{x}_2\mu+\lambda} - \frac{k^2(\rho)}{(n-2)(n-1)}K^{\mathrm{off}}, \qquad (4.27)
\end{aligned}
$$

since $c > \alpha K^{\mathrm{off}}$.

We next consider the switch on cost associated with policy $R_2$, for which we derive an upper bound. We can use a similar route as in the proof of the Claim, with the provision that the probability of reaching state $(0, n-1)$ from $(1, n-2)$, without a coupling having taken place, has to be taken into account. This probability is less or equal to $a(n-2)$ and so, again by Lemma 4.6.1

$$
\begin{aligned}
&K^{\mathrm{on}}\,\mathsf{E}^*_x\Big\{\int_0^{J_T-J_1} e^{-\alpha t}\lambda\cdot \mathbb{1}_{\{(0,n-1)\}}(X_t(R_2))dt\,\mathbb{1}_{\{T=\tau\}}\,\Big|T>1\Big\} \\
&\leq K^{\mathrm{on}}\cdot a(n-2)\frac{\lambda}{\alpha+(n-1)\mu+\lambda}\Big\{1+\sum_{l\geq 1}\Big(a(n-2)\frac{(n-1)\mu}{\alpha+(n-1)\mu+\lambda}\Big)^l\Big\} \\
&= K^{\mathrm{on}}\frac{a(n-2)\lambda}{\alpha+(n-1)\mu(1-a(n-2))+\lambda} \\
&\leq K^{\mathrm{on}}\frac{k(\rho)\lambda}{(n-2)(\alpha+(n-1)\mu(1-k(\rho)/(n-2))+\lambda)}, \qquad (4.28)
\end{aligned}
$$

for $n > 2 + k(\rho)$.

Putting Eqs. (4.27) and (4.28) together, yields the desired non-negativity of Eq. (4.20) provided $n \geq n_0(\alpha)$. A quick motivation is that Eq. (4.28) and the second term in Eq. (4.27) are of order $1/n^2$ and the first term in Eq. (4.27) is positive for large $n$ and of order $1/n$. Clearly, this can be made positive for $n$ large. A check of the choice of $n_0(\alpha)$ ensures this to be valid. One of the bounds used is that

$$
\begin{aligned}
\Gamma(\alpha)\cdot p &= \mathsf{E}_{x^*}\big\{e^{-\alpha(J_{m_0}-J_1)}\mathbb{1}_{\{J_{m_0}<J_T\}}\,|\,T>1\big\} \\
&\geq \Big(\frac{\bar{x}_2\mu}{\alpha+\bar{x}_2\mu+\lambda}\Big)^{n-2-\bar{x}_2} \\
&\geq \Big(\frac{\mu}{\alpha+\mu+\lambda}\Big)^{n_1(\alpha)} =: \frac{1}{\eta(\alpha)},
\end{aligned}
$$

for $n > n_1(\alpha)$ (cf. Eq. (4.1)). Thus, turning off is the optimal decision in $(1, n-1)$, for $n \geq n_0(\alpha)$. This proves assertion (1,ii,b).

Finally, we will prove assertion (1,iii). Therefore, assume that $c \leq \alpha K^{\mathrm{off}}$. By virtue of assertion (1,ii,a), it is optimal to idle in state $x \in B$. Let $x \in \mathbf{S}^* \setminus B$. We consider the $\xi$-restricted MDP on $\Delta(\xi)$, with $\xi = x_1 + x_2$.

Let $\delta$ be an $\alpha$-discount optimal policy on $\Delta(\xi)$. The policy $R_1$ is stationary, deterministic, say it equals $\delta^1$ that we will define next. Put $\delta^1_x = d_1 = \text{OFF}$, and $\delta^1_y = \delta_y$, $y \neq x$. The policy $R_2$ has $d_2 = \text{IDLE} = D_{X_{J_0}(R_2)}$. Put $R_{2a}$ equal to $\delta^1$. Put $R_{2d}$ the following non-stationary deterministic policy: if $X_{J_1}(R_{2d}), \ldots, X_{J_m}(R_{2d}) \notin B$, then put $D_{X_{J_m}(R_{2d})} = \delta^1_{X_{J_m}(R_{2d})-e_1}$. If $X_{J_l}(R_{2a}) \in B$ for some $l \leq m$, then put $D_{X_{J_m}(R_{2d})} = \delta^1_{X_{J_m}(R_{2d})}$.

In words, policy $R_2$ 'imitates' the decisions of $R_1$, till they couple. Coupling either takes place at time $J_1$, if an arrival occurred, or at the time that $B$ is hit by the process controlled by $R_2$. We get

$$\begin{aligned}
&v^{\alpha}_{x,\sigma_x}(\xi, R_1) - v^{\alpha}_{x,\sigma_x}(\xi, R_2) = \\
=&\mathsf{E}^*_x\Big\{ \int_0^{J_1} e^{-\alpha t} x_2 \mu K^{\text{off}} dt \Big\} \\
+&\frac{x_2\mu}{x_2\mu + \lambda} \mathsf{E}^*_x \Big\{ e^{-\alpha J_1} \int_0^{J_T - J_1} e^{-\alpha t} (\lambda K^{\text{on}} \mathbb{1}_{\{X_{t,1}(R_1)=0<X_{t,1}(R_2)\}} - c) dt \Big| T > 1 \Big\} \\
\geq& \frac{x_2\mu}{\alpha + x_2\mu + \lambda} \Big[ K^{\text{off}} - \frac{c}{\alpha} + \frac{c}{\alpha} \mathsf{E}^*_x \Big\{ e^{-\alpha(J_T - J_1)} \,\Big|\, T > 1 \Big\} \Big] > 0.
\end{aligned}$$

This concludes the proof of (1).

We will prove (2). Assume any condition (2,i) or (2,ii), say (2,ii). The other case is proved analogously. Then there exists $\alpha_0$, such that the corresponding condition (1,ii) is satisfied for $\alpha < \alpha_0$. Note that $\alpha \mapsto n_i(\alpha)$ is continuous increasing, and $n_i(\alpha) \downarrow n_i(0)$, $\alpha \downarrow 0$, $i = 0, 1$. Let $\delta^\alpha$ denote an $\alpha$-discount optimal policy.

First of all, for $x \in B$, $\delta^\alpha_x = \text{IDLE}$, $\alpha \leq \alpha_0$. There exists $\alpha_1 \leq \alpha_0$, such that $\lfloor n_0(\alpha) \rfloor = \lfloor n_0(0) \rfloor$. For $\alpha \leq \alpha_1$ we have that

$$\mathbf{S}^* \setminus \{\Delta(n_0(0)) \cup B\} = \mathbf{S}^* \setminus \{\Delta(n_0(\alpha)) \cup B\}. \tag{4.29}$$

Hence $\delta^\alpha_x = \text{OFF}$, for $x \in \mathbf{S}^* \setminus \{\Delta(n_0(0)) \cup B\}$, $\alpha \leq \alpha_1$.

Theorem 4.3.3 implies for any sequence $\alpha_m \to 0$ that any limit point of $\{\delta^{\alpha_m}\}_m$ is average optimal. Thus there exists an average optimal policy that idles in $B$ and switches off in $\mathbf{S}^* \setminus \{\Delta(n_0(0)) \cup B\}$. In fact, since in the average cost case it is sufficient to solve a total cost problem on $\Delta(n)$ (cf. Lemma 4.4.1), we could have repeated all the arguments leading to assertion (1), for $\alpha = 0$.

We will finally show (3), that there exists a strong Blackwell optimal policy. From the proof of (2) Eq. (4.29), IDLE is optimal in $B$, and OFF in $\mathbf{S}^* \setminus \{\Delta(n_0(0)) \cup B\}$, for all $\alpha \leq \alpha_1$. So only within $\Delta(n_0(0)$ the optimal policy may change as a function of $\alpha$.

By virtue of Lemma 4.4.1, for determining the decisions under an $\alpha$-discount optimal policy, or an average cost optimal policy on the set of states $\Delta(n_0(0))$, it is sufficient solve the problem as a total $\alpha$-discounted or total expected cost problem, where we take $(0, \lfloor n_0(0) \rfloor)$ as a zero cost absorbing state. The total expected cost problem is equivalent to a unichain average cost problem, with absorbing state $(0, \lfloor n_0(0) \rfloor)$ that is reached with probability 1 in finite expected time under any policy. This is a finite state and action, continuous time MDP. Since it is uniformisable, one solve the equivalent uniformised, discrete time MDP. The required conditions from [26, Chapter 3, Corollary 1] are met and so there exists a strong Blackwell optimal policy on $\Delta(n_0(0))$. □

**Lemma 4.4.4.** *Algorithm 3 converges in a finite number of steps, and computes an $\alpha$-discount optimal policy, for any $\alpha > 0$. It computes an average optimal policy by taking $\alpha = 0$ in the algorithm.*

*Proof.* The algorithm has to perform at most $n_0(\alpha)$ iterations, and thus terminates in finitely many steps.

On $\Delta(1)$, it is easily checked that the algorithm computes $v^\alpha(1)$, the $\alpha$-discount value function, $\alpha > 0$, for the 1-perturbed problem. It generates an optimal policy on $\Delta(1) \setminus \{(0,1)\}$.

Consider iteration $n$, concerning the states of $\Delta(n)$. Assume that iteration $n-1$ has yielded the values $v^\alpha_x(n-1)$ for $x \in \Delta(n) \setminus \{(0, n-1)\}$, and an SC optimal policy.

As in the proof of Lemma 4.4.1 we may deduce for $x \in \Delta(n-1) \setminus \{(0, n-1)\}$ that

$$v^\alpha_x(n) = v^\alpha_x(n-1) + \prod_{l=x_2}^{n-2} t_k \cdot v^\alpha_{(0,n-1)}(n). \tag{4.30}$$

However, $v^\alpha_{(0,n-1)}(n)$ has to be computed. Under the conditions on $c$, $K^{\mathrm{off}}$ and $K^{\mathrm{on}}$, idling is optimal in $(0, n-1)$ for both the restricted and unrestricted problems, by Theorem 4.2.2. Use the optimality equation (4.13) for the $n$-restricted problem for state $(0, n-1)$, where we plug in the decision IDLE. Then, the only possible transition is to state $(1, n-2)$, since the transition to $(0, n)$ gives no contribution to the cost. Using Eq. (4.30) for $x = (1, n-2)$, we obtain an equation for state $(0, n-1)$ with one unknown, the solution of which is given by the algorithm, step 1.

This yields the new values $v^\alpha_x(n)$ explicitly for $x \in \Delta(n-1)$.

Lemma 4.2.3 allows to deduce that $v^\alpha(n-1)$ is and $\mathcal{S}uper$ and $\mathcal{S}uper\mathcal{C}$ on $\Delta(n-1)$, as is the yet unknown $v^\alpha(n)$ on $\Delta(n)$. This implies that there is an SC $\alpha$-discounted optimal extension of the SC optimal policy on $\Delta(n-1)$ to

$\Delta(n)$. For the specified values of the SC optimal policy on $\Delta(n-1)$ in the algorithm, this implies that the optimal decision in the specified states of $A(n)$ is known. Thus, it reduces the optimisation problem. The choice of how to resolve ties is then somewhat arbitrary, but systematic.

Since $v^{\alpha}(n)$ is the unique solution of the optimality equation (4.13) on $\Delta(n)$, we can then solve the missing values $v_x^{\alpha}(n)$, $x \in A(n)$, from Eq. (4.13).

In the average cost case, the arguments are analogous. In iteration $n$ the minimum total expected cost $w_x(n)$, $x \in \Delta(n) \setminus \{(0,n)\}$, till absorption in $(0,n)$ is computed. However, it is denoted by $v^0(n)$ in the algorithm. Note that $\prod_{l=x_2}^{n-2} t_k = 1$, if $\alpha = 0$. Thus Eq. (4.30) reduces to

$$v_x^0(n) = w_x(n) = w_x(n-1) + w_{(0,n-1)}(n) = v_x^0(n-1) + v_{(0,n-1)}^0(n),$$

which is the basis for extending the solution on $\Delta(n-1)$ to a solution on $\Delta(n)$ in the algorithm. Clearly, the validity of this equation can be deduced by similar arguments that have been used in the proof of Lemma 4.4.1. $\square$

## 4.5 Propagation results

Recall the value iteration scheme for $x \in \mathbf{S}, 0 < \bar{\alpha} < 1$,

$$\bar{v}_x^{\bar{\alpha},0}(N) = 0,$$

and then iterating for $n = 0, 1, \ldots$

$$\begin{aligned}\bar{v}_x^{\bar{\alpha},n+1}(N) = {} & x_1\bar{c} + (1-\bar{\alpha})\Big(\lambda\big\{\mathbb{1}_{\{x_1>0\}}\bar{v}_{x-e_1+e_2}^{\bar{\alpha},n}(N) + \mathbb{1}_{\{x_1=0\}}(\bar{v}_{x+e_2}^{\bar{\alpha},n}(N) + K^{\mathrm{on}})\big\} \\ & + \{x_2 \wedge N\}\mu \min\{\bar{v}_{x+e_1-e_2}^{\bar{\alpha},n}(N), \bar{v}_{x-e_2}^{\bar{\alpha},n}(N) + K^{\mathrm{off}}\} + (N-x_2)^+\mu\bar{v}_x^{\bar{\alpha},n}(N)\Big).\end{aligned}$$

Now we introduce the following five operators

$$\begin{aligned}\mathcal{T}_A f_x := {} & \mathbb{1}_{\{x_1>0\}} f_{x-e_1+e_2} + \mathbb{1}_{\{x_1=0\}}(f_{x+e_2} + K^{\mathrm{on}}); \\ \mathcal{T}_{I/O}^N f_x := {} & \min\{\frac{x_2}{N}, 1\} \min\{f_{x+e_1-e_2}, f_{x-e_2} + K^{\mathrm{off}}\} + (1 - \frac{x_2}{N})^+ f_x; \\ \mathcal{T}_C f_x := {} & x_1\bar{c} + f(x),\end{aligned}$$

for $x \in \mathbf{S}$. Further,

$$\begin{aligned}\mathcal{T}_{UNIF}^N(f^1, f^2) := {} & \lambda f^1 + N\mu f^2; \\ \mathcal{T}_{DISC}^{\bar{\alpha}}(f) := {} & (1-\bar{\alpha})f.\end{aligned}$$

This implies that

$$\bar{v}_x^{\bar{\alpha},n+1}(N) = \mathcal{T}_C(\mathcal{T}_{DISC}^{\bar{\alpha}}(\mathcal{T}_{UNIF}^{N}(\mathcal{T}_A, \mathcal{T}_{I/O}^{N})))\bar{v}_x^{\bar{\alpha},n}(N). \tag{4.31}$$

The following short-hand notation is convenient

$$\begin{aligned}
\mathcal{S}uper[f_x] &:= f_{x+e_1+e_2} - f_{x+e_1} - f_{x+e_2} + f_x;\\
\mathcal{S}uper\mathcal{C}[f_x] &:= f_{x+2e_1} - f_{x+e_1} - f_{x+e_1+e_2} + f_{x+e_2};\\
\mathcal{D}iff_{K^{\mathrm{on}}}[f_x] &:= f_{x+e_1} - f_x + K^{\mathrm{on}};
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{S}uper &:= \{f : \mathcal{S}uper[f_x] \geq 0,\ \text{for } x_1, x_2 \geq 0\} && \text{(supermodularity)};\\
\mathcal{S}uper\mathcal{C} &:= \{f : \mathcal{S}uper\mathcal{C}[f_x] \geq 0,\ \text{for } x_1, x_2 \geq 0\} && \text{(superconvexity)};\\
\mathcal{D}iff_{K^{\mathrm{on}}} &:= \{f : \mathcal{D}iff_{K^{\mathrm{on}}}[f_x] \geq 0,\ \text{for } x_1, x_2 \geq 0\} && \text{(bounded difference)}.
\end{aligned}$$

The following lemma gives all the propagation results. Clearly, there is some overlap with [1]. Propagations 1 and 2 appear in an identical form. Propagations 4 and 5 are similar as in [1], except for the extra term $K^{\mathrm{off}}$. Propagations 3 and 6 are also present there, but we treat them in a more direct manner. Propagations 7-15 are absent in [1]. For completeness we provide all proofs in detail here.

**Lemma 4.5.1.** *The following fifteen propagations hold*

$$\begin{aligned}
\mathcal{T}_A &: \mathcal{S}uper \xrightarrow{1} \mathcal{S}uper, \mathcal{S}uper\mathcal{C} \cap \mathcal{D}iff_{K^{\mathrm{on}}} \xrightarrow{2} \mathcal{S}uper\mathcal{C}, \mathcal{D}iff_{K^{\mathrm{on}}} \xrightarrow{3} \mathcal{D}iff_{K^{\mathrm{on}}};\\
\mathcal{T}_{I/O}^{N} &: \mathcal{S}uper \cap \mathcal{S}uper\mathcal{C} \xrightarrow{4,5} \mathcal{S}uper \cap \mathcal{S}uper\mathcal{C},\ \mathcal{D}iff_{K^{\mathrm{on}}} \xrightarrow{6} \mathcal{D}iff_{K^{\mathrm{on}}};\\
\mathcal{T}_{UNIF}^{N} &: \mathcal{S}uper \xrightarrow{7} \mathcal{S}uper,\ \mathcal{S}uper\mathcal{C} \xrightarrow{8} \mathcal{S}uper\mathcal{C},\ \mathcal{D}iff_{K^{\mathrm{on}}} \xrightarrow{9} \mathcal{D}iff_{K^{\mathrm{on}}};\\
\mathcal{T}_{DISC}^{\bar{\alpha}} &: \mathcal{S}uper \xrightarrow{10} \mathcal{S}uper,\ \mathcal{S}uper\mathcal{C} \xrightarrow{11} \mathcal{S}uper\mathcal{C},\ \mathcal{D}iff_{K^{\mathrm{on}}} \xrightarrow{12} \mathcal{D}iff_{K^{\mathrm{on}}};\\
\mathcal{T}_C &: \mathcal{S}uper \xrightarrow{13} \mathcal{S}uper,\ \mathcal{S}uper\mathcal{C} \xrightarrow{14} \mathcal{S}uper\mathcal{C},\ \mathcal{D}iff_{K^{\mathrm{on}}} \xrightarrow{15} \mathcal{D}iff_{K^{\mathrm{on}}}.
\end{aligned}$$

*Proof.* Let $x \in \mathbf{S}$.
*Proof of 1.* Suppose that $f \in \mathcal{S}uper$, then

$$\begin{aligned}
&\mathcal{T}_A f_{x+e_1+e_2} - \mathcal{T}_A f_{x+e_1} - \mathcal{T}_A f_{x+e_2} + \mathcal{T}_A f_x\\
&= \mathbb{1}_{\{x_1>0\}} \mathcal{S}uper[f_{x-e_1+e_2}]\\
&\geq 0.
\end{aligned}$$

*Proof of 2.* Suppose that $f \in \mathcal{S}uper\mathcal{C} \cap \mathcal{D}iff_{K^{\text{on}}}$, then

$$\begin{aligned}\mathcal{T}_A f_{x+2e_1} - \mathcal{T}_A f_{x+e_1} - \mathcal{T}_A f_{x+e_1+e_2} + \mathcal{T}_A f_{x+e_2} \\ &= \mathbb{1}_{\{x_1>0\}} \mathcal{S}uper\mathcal{C}[f_{x-e_1+e_2}] + \mathbb{1}_{\{x_1=0\}} \mathcal{D}iff_{K^{\text{on}}}[f_{x+e_2}] \\ &\geq 0.\end{aligned}$$

*Proof of 3.* Suppose that $f \in \mathcal{D}iff_{K^{\text{on}}}$, then

$$\begin{aligned}\mathcal{T}_A f_{x+e_1} - \mathcal{T}_A f_x + K^{\text{on}} \\ &= \mathbb{1}_{\{x_1>0\}} \mathcal{D}iff_{K^{\text{on}}}[f_{x-e_1+e_2}] \\ &\geq 0.\end{aligned}$$

*Proof of 4.* Suppose $f \in \mathcal{S}uper \cap \mathcal{S}uper\mathcal{C}$. Then for $x_2 < N$

$$\begin{aligned}N\big(\mathcal{T}^N_{I/O} f_{x+e_1+e_2} - \mathcal{T}^N_{I/O} f_{x+e_1} - \mathcal{T}^N_{I/O} f_{x+e_2} + \mathcal{T}^N_{I/O} f_x\big) \\ =&(x_2+1)\min\{f_{x+2e_1}, f_{x+e_1}+K^{\text{off}}\} + (N-x_2-1) f_{x+e_1+e_2} \\ &- x_2 \min\{f_{x+2e_1-e_2}, f_{x+e_1-e_2}+K^{\text{off}}\} - (N-x_2) f_{x+e_1} \\ &- (x_2+1)\min\{f_{x+e_1}, f_x+K^{\text{off}}\} - (N-x_2-1) f_{x+e_2} \\ &+ x_2 \min\{f_{x+e_1-e_2}, f_{x-e_2}+K^{\text{off}}\} + (N-x_2) f_x. \end{aligned} \tag{4.32}$$

Case (I): $\pi_{x+e_1+e_2} = \pi_x = \text{OFF}$.
Then by Eq. (4.32)

$$\begin{aligned}N\big(\mathcal{T}^N_{I/O} f_{x+e_1+e_2} - \mathcal{T}^N_{I/O} f_{x+e_1} - \mathcal{T}^N_{I/O} f_{x+e_2} + \mathcal{T}^N_{I/O} f_x\big) \\ &\geq x_2 \mathcal{S}uper[f_{x-e_2}] + f_{x+e_1} - f_x \\ &\quad + (N-x_2-1)\mathcal{S}uper[f_x] - f_{x+e_1} + f_x \\ &\geq 0.\end{aligned}$$

Case (II): $\pi_{x+e_1+e_2} = \pi_x = \text{IDLE}$.
Then by Eq. (4.32)

$$\begin{aligned}N\big(\mathcal{T}^N_{I/O} f_{x+e_1+e_2} - \mathcal{T}^N_{I/O} f_{x+e_1} - \mathcal{T}^N_{I/O} f_{x+e_2} + \mathcal{T}^N_{I/O} f_x\big) \\ &\geq x_2 \mathcal{S}uper[f_{x+e_1-e_2}] + f_{x+e_2} - f_{x+e_1} \\ &\quad + (N-x_2-1)\mathcal{S}uper[f_x] - f_{x+e_1} + f_x \\ &\geq f_{x+2e_1} - 2f_{x+e_1} + f_x \\ &\geq 0,\end{aligned}$$

where for the last inequality we have used convexity.

Case (III) $\pi_{x+e_1+e_2} = \textsc{off}$, $\pi_x = \textsc{idle}$.
Then by Eq. (4.32)

$$
\begin{aligned}
&N\big(\mathcal{T}_{I/O}^{N} f_{x+e_1+e_2} - \mathcal{T}_{I/O}^{N} f_{x+e_1} - \mathcal{T}_{I/O}^{N} f_{x+e_2} + \mathcal{T}_{I/O}^{N} f_x\big) \\
&\qquad \geq x_2[f_{x+e_1} - f_{x+e_1-e_2} - f_{x+e_1} + f_{x+e_1-e_2}] + f_{x+e_1} - f_x \\
&\qquad\quad + (N - x_2 - 1)\mathcal{S}uper[f_x] - f_{x+e_1} + f_x \\
&\qquad \geq 0.
\end{aligned}
$$

There is no fourth case due to $f \in \mathcal{S}uper \cap \mathcal{S}uper\mathcal{C}$. If $x_2 \geq N$, then the operator $\mathcal{T}_{I/O}^{N}$ reduces to $\mathcal{T}_{I/O}^{N} f_x = \min\{f_{x+e_1-e_2}, f_{x-e_2} + K^{\text{off}}\}$. The subsequent propagation follows as a simpler case of the above one.

*Proof of 5.* Suppose that $f \in \mathcal{S}uper \cap \mathcal{S}uper\mathcal{C}$. Then for $x_2 < N$

$$
\begin{aligned}
&N\big(\mathcal{T}_{I/O}^{N} f_{x+2e_1} - \mathcal{T}_{I/O}^{N} f_{x+e_1} - \mathcal{T}_{I/O}^{N} f_{x+e_1+e_2} + \mathcal{T}_{I/O}^{N} f_{x+e_2}\big) \\
&\quad = x_2 \min\{f_{x+3e_1-e_2}, f_{x+2e_1-e_2} + K^{\text{off}}\} + (N - x_2)_{x+2e_1} \\
&\qquad - x_2 \min\{f_{x+2e_1-e_2}, f_{x+e_1-e_2} + K^{\text{off}}\} - (N - x_2) f_{x+e_1} \\
&\qquad - (x_2 + 1) \min\{f_{x+2e_1}, f_{x+e_1} + K^{\text{off}}\} - (N - x_2 - 1) f_{x+e_1+e_2} \\
&\qquad + (x_2 + 1) \min\{f_{x+e_1}, f_x + K^{\text{off}}\} + (N - x_2 - 1) f_{x+e_2}.
\end{aligned}
\tag{4.33}
$$

Case (I) $\pi_{x+2e_1} = \pi_{x+e_2} = \textsc{off}$. Then by Eq. (4.33)

$$
\begin{aligned}
&N\big(\mathcal{T}_{I/O}^{N} f_{x+2e_1} - \mathcal{T}_{I/O}^{N} f_{x+e_1} - \mathcal{T}_{I/O}^{N} f_{x+e_1+e_2} + \mathcal{T}_{I/O}^{N} f_{x+e_2}\big) \\
&\qquad \geq x_2 \mathcal{S}uper\mathcal{C}[f_{x-e_2}] - f_{x+e_1} + f_x \\
&\qquad\quad + (N - x_2 - 1)\mathcal{S}uper\mathcal{C}[f_x] + f_{x+2e_1} - f_{x+e_1} \\
&\qquad \geq f_{x+2e_1} - 2 f_{x+e_1} + f_x \\
&\qquad \geq 0,
\end{aligned}
$$

where convexity is used for the last inequality.

Case (II) $\pi_{x+2e_1} = \pi_{x+e_2} = \textsc{idle}$.
Then by Eq. (4.33)

$$
\begin{aligned}
&N\big(\mathcal{T}_{I/O}^{N} f_{x+2e_1} - \mathcal{T}_{I/O}^{N} f_{x+e_1} - \mathcal{T}_{I/O}^{N} f_{x+e_1+e_2} + \mathcal{T}_{I/O}^{N} f_{x+e_2}\big) \\
&\qquad \geq x_2 \mathcal{S}uper\mathcal{C}[f_{x+e_1-e_2}] - f_{x+2e_1} + f_{x+e_1} \\
&\qquad\quad + (N - x_2 - 1)\mathcal{S}uper\mathcal{C}[f_x] + f_{x+2e_1} - f_{x+e_1} \\
&\qquad \geq 0.
\end{aligned}
$$

Case (III) $\pi_{x+2e_1} = \text{OFF}$, $\pi_{x+e_2} = \text{IDLE}$. Then by Eq. (4.33)

$$\begin{aligned}
N\big(&\mathcal{T}^N_{I/O} f_{x+2e_1} - \mathcal{T}^N_{I/O} f_{x+e_1} - \mathcal{T}^N_{I/O} f_{x+e_1+e_2} + \mathcal{T}^N_{I/O} f_{x+e_2}\big)\\
&\geq x_2[f_{x+2e_1-e_2} - f_{x+2e_1-e_2} - f_{x+e_1} + f_{x+e_1}] + f_{x+2e_1} - f_{x+e_1}\\
&\qquad + (N - x_2 - 1)\mathcal{S}uper\mathcal{C}[f_x] - f_{x+2e_1} + f_{x+e_1}\\
&\geq 0.
\end{aligned}$$

Again, the propagation for $x_2 \geq N$ is straightforward.

*Proof of 6.* Suppose that $f \in \mathcal{D}iff_{K^{\text{on}}}$, then

$$\begin{aligned}
&\mathcal{T}^N_{I/O} f_{x+e_1} - \mathcal{T}^N_{I/O} f_x + K^{\text{on}}\\
&\qquad = \min\{\frac{x_2}{N}, 1\} \min\{f_{x+2e_1-e_2}, f_{x+e_1-e_2} + K^{\text{off}}\} + (1 - \frac{x_2}{N})^+ f_{x+e_1}\\
&\qquad\quad - \min\{\frac{x_2}{N}, 1\} \min\{f_{x+e_1-e_2}, f_{x-e_2} + K^{\text{off}}\} - (1 - \frac{x_2}{N})^+ f_x.
\end{aligned} \tag{4.34}$$

Case (I) $\pi_{x+e_1} = \text{OFF}$.
By Eq. (4.34)

$$\begin{aligned}
&\mathcal{T}^N_{I/O} f_{x+e_1} - \mathcal{T}^N_{I/O} f_x + K^{\text{on}}\\
&\qquad \geq \min\{\frac{x_2}{N}, 1\}\mathcal{D}iff_{K^{\text{on}}}[f_{x-e_2}] + (1 - \frac{x_2}{N})^+ \mathcal{D}iff_{K^{\text{on}}}[f_x]\\
&\qquad \geq 0.
\end{aligned}$$

Case (II) $\pi_{x+e_1} = \text{IDLE}$.
By Eq. (4.34)

$$\begin{aligned}
&\mathcal{T}^N_{I/O} f_{x+e_1} - \mathcal{T}^N_{I/O} f_x + K^{\text{on}}\\
&\qquad \geq \min\{\frac{x_2}{N}, 1\}\mathcal{D}iff_{K^{\text{on}}}[f_{x+e_1-e_2}] + (1 - \frac{x_2}{N})^+ \mathcal{D}iff_{K^{\text{on}}}[f_x]\\
&\qquad \geq 0.
\end{aligned}$$

*Proof of 7.* Suppose that $f^1, f^2 \in \mathcal{S}uper$, then

$$\begin{aligned}
&\mathcal{T}^N_{UNIF}(f^1, f^2)_{x+e_1+e_2} - \mathcal{T}^N_{UNIF}(f^1, f^2)_{x+e_1}\\
&\qquad\quad - \mathcal{T}^N_{UNIF}(f^1, f^2)_{x+e_2} + \mathcal{T}^N_{UNIF}(f^1, f^2)_x\\
&\qquad = \lambda \mathcal{S}uper[f^1_x] + N\mu \mathcal{S}uper[f^2_x]\\
&\qquad \geq 0.
\end{aligned}$$

*Proof of 8.* Suppose that $f^1, f^2 \in \mathcal{S}uper\mathcal{C}$, then

$$\begin{aligned}
&\mathcal{T}^N_{UNIF}(f^1, f^2)_{x+2e_1} - \mathcal{T}^N_{UNIF}(f^1, f^2)_{x+e_1} \\
&\qquad - \mathcal{T}^N_{UNIF}(f^1, f^2)_{x+e_1+e_2} + \mathcal{T}^N_{UNIF}(f^1, f^2)_{x+e_2} \\
&\quad = \lambda \mathcal{S}uper\mathcal{C}[f^1_x] + N\mu \mathcal{S}uper\mathcal{C}[f^2_x] \\
&\quad \geq 0.
\end{aligned}$$

*Proof of 9.* Suppose that $f^1$, $f^2 \in \mathcal{D}iff_{K^{\text{on}}}$. Notice that $\lambda + N\mu = 1$, so

$$\begin{aligned}
&\mathcal{T}^N_{UNIF}(f^1, f^2)_{x+e_1} - \mathcal{T}^N_{UNIF}(f^1, f^2)_x + K^{\text{on}} \\
&\quad = \lambda \mathcal{D}iff_{K^{\text{on}}}[f^1_x] + N\mu \mathcal{D}iff_{K^{\text{on}}}[f^2_x] \\
&\quad \geq 0.
\end{aligned}$$

*Proof of 10.* Suppose that $f \in \mathcal{S}uper$, then

$$\begin{aligned}
&\mathcal{T}^{\bar{\alpha}}_{DISC} f_{x+e_1+e_2} - \mathcal{T}^{\bar{\alpha}}_{DISC} f_{x+e_1} - \mathcal{T}^{\bar{\alpha}}_{DISC} f_{x+e_2} + \mathcal{T}^{\bar{\alpha}}_{DISC} f_x \\
&\quad = (1 - \bar{\alpha}) \mathcal{S}uper[f_x] \\
&\quad \geq 0.
\end{aligned}$$

*Proof of 11.* Suppose that $f \in \mathcal{S}uper\mathcal{C}$, then

$$\begin{aligned}
&\mathcal{T}^{\bar{\alpha}}_{DISC} f_{x+2e_1} - \mathcal{T}^{\bar{\alpha}}_{DISC} f_{x+e_1} - \mathcal{T}^{\bar{\alpha}}_{DISC} f_{x+e_1+e_2} + \mathcal{T}^{\bar{\alpha}}_{DISC} f_{x+e_2} \\
&\quad = (1 - \bar{\alpha}) \mathcal{S}uper\mathcal{C}[f_x] \\
&\quad \geq 0.
\end{aligned}$$

*Proof of 12.* Suppose that $f \in \mathcal{D}iff_{K^{\text{on}}}$, then

$$\begin{aligned}
&\mathcal{T}^{\bar{\alpha}}_{DISC} f_{x+e_1} - \mathcal{T}^{\bar{\alpha}}_{DISC} f_x + K^{\text{on}} \\
&\quad = (1 - \bar{\alpha}) \mathcal{D}iff_{K^{\text{on}}}[f_x] + \alpha K^{\text{on}} \\
&\quad \geq 0.
\end{aligned}$$

*Proof of 13.* Suppose that $f \in \mathcal{S}uper$, then

$$\begin{aligned}
&\mathcal{T}_C f_{x+e_1+e_2} - \mathcal{T}_C f_{x+e_1} - \mathcal{T}_C f_{x+e_2} + \mathcal{T}_C f_x \\
&\quad = (x_1 + 1)\bar{c} - (x_1 + 1)\bar{c} - x_1 \bar{c} + x_1 \bar{c} + \mathcal{S}uper[f_x] \\
&\quad \geq 0.
\end{aligned}$$

*Proof of 14.* Suppose that $f \in \mathcal{S}uper\mathcal{C}$, then

$$\begin{aligned}
&\mathcal{T}_C f_{x+2e_1} - \mathcal{T}_C f_{x+e_1} - \mathcal{T}_C f_{x+e_1+e_2} + \mathcal{T}_C f_{x+e_2} \\
&\quad = (x_1+2)\bar{c} - (x_1+1)\bar{c} - (x_1+1)\bar{c} + x_1\bar{c} + \mathcal{S}uper\mathcal{C}[f_x] \\
&\quad \geq 0.
\end{aligned}$$

*Proof of 15.* Suppose that $f \in \mathcal{D}iff_{K^{\mathrm{on}}}$, then

$$\begin{aligned}
&\mathcal{T}_C f_{x+e_1} - \mathcal{T}_C f_x + K^{\mathrm{on}} \\
&\quad = (x_1+1)\bar{c} - x_1\bar{c} + \mathcal{D}iff_{K^{\mathrm{on}}}[f_x] \\
&\quad \geq 0.
\end{aligned}$$

□

Lemma 4.5.1 allows to finish the proof of Theorem 4.3.2.

*Proof of Theorem 4.3.2.* Put $\bar{v}_x^{\bar{\alpha},0}(N) \equiv 0$. Then $\bar{v}_x^{\bar{\alpha},0}(N) \in \mathcal{S}uper \cap \mathcal{S}uper\mathcal{C} \cap \mathcal{D}iff_{K^{\mathrm{on}}}$. VI is equivalent to the map given in Eq. (4.31) for $n \geq 0$. Lemma 4.5.1 implies that

$$\begin{aligned}
\mathcal{T}_C(\mathcal{T}^{\bar{\alpha}}_{DISC}(\mathcal{T}^N_{UNIF}(\mathcal{T}_A, \mathcal{T}^N_{I/O}))) : \mathcal{S}uper \cap \mathcal{S}uper\mathcal{C} \cap \mathcal{D}iff_{K^{\mathrm{on}}}& \\
\to \mathcal{S}uper \cap \mathcal{S}uper\mathcal{C} \cap \mathcal{D}iff_{K^{\mathrm{on}}}&.
\end{aligned}$$

In other words

$$\bar{v}_x^{\bar{\alpha},1}(N) = \mathcal{T}_C(\mathcal{T}^{\bar{\alpha}}_{DISC}(\mathcal{T}^N_{UNIF}(\mathcal{T}_A, \mathcal{T}^N_{I/O})))\bar{v}_x^{\bar{\alpha},0}(N) \in \mathcal{S}uper \cap \mathcal{S}uper\mathcal{C} \cap \mathcal{D}iff_{K^{\mathrm{on}}}.$$

By induction we have that $\bar{v}_x^{\bar{\alpha},n}(N) \in \mathcal{S}uper \cap \mathcal{S}uper\mathcal{C} \cap \mathcal{D}iff_{K^{\mathrm{on}}}$ for all $n \geq 0$. □

## 4.6 Probability bounds

Consider the M/M/∞-queue $\{X_{t,2}\}_t$, where we use the notation of the previous sections. Recall that the arrival rate is $\lambda$ and the departure rate $\mu n$ if $n$ jobs are present. We denote the successive jump times by $J_0 = 0, J_1, \ldots,$ and assume that the trajectories are right-continuous. Recall the definition of $a(n)$ in Eq. (4.22) as the probability that state $n+1$ is reached before 0, given a start in state $n$. Then,

$$a(0) = 0, \quad \text{and } a(1) = \frac{\lambda}{\lambda + \mu}.$$

Denote $\rho = \lambda/\mu$.

**Lemma 4.6.1.** *It holds that*

$$a(n) \leq \frac{(1+\rho)^2 e^{\rho}}{n\rho}.$$

*Proof.* To estimate $a(n)$, we will use a recursion on $f^n(i)$, the probability that there are $n$ jobs in the system before the system is empty, given a start in state $i$. Denote $\Delta f^n(i) := f^n(i+1) - f^n(i)$. Then $f^n(0) = 0$, $f^n(n) = 1$ and further we have

$$f^n(i) = \frac{\lambda}{\lambda + i\mu} f^n(i+1) + \frac{i\mu}{\lambda + i\mu} f^n(i-1).$$

This implies

$$\lambda \cdot \Delta f^n(i) = i\mu \cdot \Delta f^n(i-1),$$

and thus,

$$\Delta f^n(i) = \rho^{-1} i \cdot \Delta f^n(i-1) = \rho^{-i} i! \cdot \Delta f^n(0).$$

This leads to

$$f^n(i) = \sum_{k=0}^{i-1} \Delta f^n(k) = \sum_{k=0}^{i-1} \rho^{-k} k! \cdot \Delta f^n(0) = \sum_{k=0}^{i-1} \rho^{-k} k! \cdot f^n(1).$$

Now we use that $f^n(n) = 1$. This yields,

$$1 = f^n(n) = \sum_{k=0}^{n-1} \rho^{-k} k! \cdot f^n(1).$$

Thus,

$$f^n(1) = \left( \sum_{k=0}^{n-1} \rho^{-k} k! \right)^{-1},$$

so that

$$f^n(i) = \frac{\sum_{k=0}^{i-1} \rho^{-k} k!}{\sum_{k=0}^{n-1} \rho^{-k} k!}.$$

By definition, we have that $a(n) = f^{n+1}(n)$, hence

$$a(n) = \frac{\sum_{k=0}^{n-1} \rho^{-k} k!}{\sum_{k=0}^{n} \rho^{-k} k!}, \quad n \geq 1. \tag{4.35}$$

It follows trivially that $a(n)$ meets the recursion

$$a(n) = \frac{\lambda}{n\mu(1-a(n-1))+\lambda}.$$

This easily implies that $\lim_{n\to\infty} a(n) = 0$. Moreover, the recursion allows to conclude that $a(n) > a(n-1) \iff a(n-1) < \{1 \wedge \frac{\rho}{n}\}$, and that $a(n) < a(n-1)$ implies $a(n+1) < a(n)$. Furthermore, using Eq. (4.35), it can be shown that $\rho \in [n, n+1]$ implies that $a(n+2) < a(n+1)$.

This leads to the following conclusion:

- $a(0) < a(1) \leq \cdots \leq a(\lfloor\rho\rfloor)$;
- $a(\lceil\rho\rceil) > a((\lceil\rho\rceil + 1) > a(\lceil\rho\rceil + 2) > \cdots$.

Thus,

$$a(\lfloor\rho\rfloor) \wedge a(\lceil\rho\rceil) \geq a(n), n = 0, 1, \ldots.$$

The recursion immediately implies for $n = 0, 1, \ldots$, that

$$a(n) \leq n^{-1} \frac{\rho}{(1-a(\lfloor\rho\rfloor)) \vee (1-a(\lceil\rho\rceil))}.$$

We can bound the denominator as follows. First,

$$\begin{aligned}
1 - a(\lfloor\rho\rfloor) =& \frac{\rho^{-\lfloor\rho\rfloor}\,\lfloor\rho\rfloor!}{\sum_{k=0}^{\lfloor\rho\rfloor} \rho^{-k} k!} \\
\geq& \frac{\rho^{-\lfloor\rho\rfloor}(\lfloor\rho\rfloor)!}{\lfloor\rho\rfloor + 1} \\
=& \frac{1}{\frac{\rho^{\lfloor\rho\rfloor}}{(\lfloor\rho\rfloor)!}(\lfloor\rho\rfloor + 1)} \\
\geq& \frac{e^{-\rho}}{\lfloor\rho\rfloor + 1}.
\end{aligned}$$

Similarly,

$$1 - a(\lceil\rho\rceil) \geq \frac{\rho e^{-\rho}}{\lceil\rho\rceil(1+\rho)},$$

which gives the smaller lower bound. This implies that

$$\begin{aligned}
a(n) \leq& n^{-1}\Big\{\lceil\rho\rceil \frac{1+\rho}{\rho} e^{\rho} \wedge (\lfloor\rho\rfloor + 1)e^{\rho}\Big\} \\
\leq& \frac{(1+\rho)^2}{n\rho} e^{\rho}.
\end{aligned}$$

□

# 5 Competing queues with abandonments

This chapter is based on Bhulai et al. [10], in preparation.

## 5.1 Introduction

In this chapter we consider a server assignment problem. There are $K$ classes of customers, each customer class $i$ has holding cost $c_i$, $1 \leq i \leq K$. There is a single server that can serve class $i$ with rate $\mu_i$. Arrivals occur according to independent Poisson streams. The question we address in this chapter is: what service policy minimises the expected average cost?

In the $K$-competing queues model without abandonments due to impatience it is well-known that the $c\mu$-rule is optimal. The $c\mu$-rule gives full priority to the queue with the highest index $c_i\mu_i$, the queue that gives the highest cost reduction per unit time. This result was shown to be optimal in 1985 simultaneously by Baras et al. [8] and by Buyukkoc et al. [22].

Recently there has been revived interest in the $K$-competing queues model, with the additional feature of customer abandonments due to impatience. Say that customers of type $i$ leave the system with rate $\beta_i$. In this case the $c\mu$-rule is not always optimal. When we model this problem as a continuous time Markov decision process, the impatience departures induce the transition rates to be unbounded. Hence, uniformisation is not possible and the standard (discrete time) techniques are not available. In the literature several approaches have been tried to deal with this difficulty. We may categorise them in three classes.

1. Some literature (see [5], [7] and [46]) studies a relaxation or approximate version of the original problem. The obtained policies may serve as a heuristic.

2. Another method is to use model specific coupling techniques to obtain an optimal policy. Typically, these papers (see [60] and [27]) have limitations to special cases as the coupling gets more tedious in a more general setting.

3. The third approach is to apply a truncation to make the process uniformisable. Then discrete time techniques are used to derive properties implying the optimal policy to have the desired structure (see [27], [11] and Chapters 3 and 4). This is the solution method that we will follow within this chapter.

The first approach is most prominent in the literature. Atar et al. in [5] consider a $K$-competing queues problem with many servers. In their paper the $c\mu/\beta$-rule is introduced. This rule gives priority to the queue with the highest index $c_i\mu_i/\beta_i$. They show that in the overload regime as the number of servers tends to infinity, it is asymptotically optimal to follow the $c\mu/\beta$-rule.

Ayesta et al. [7] studied the problem as well. They derive an index policy similar to the $c\mu/\beta$-rule by analytically solving the case with one or two customers present and no arrivals.

Larranaga et al. [46] have studied a fluid approximation of the multi-server variant of the competing queues problem. In this fluid approximation optimality of the $c\mu/\beta$-rule in the overload regime is shown and it is shown that for $K = 2$ a switching curve policy is optimal in the underload regime.

Other literature does not focus on heuristics, but tries to find a subset of the input parameters for which an index policy can be proven to be optimal. Salch et al. [60] studied the competing queues system with a restriction to a maximum of $K$ arrivals. With the use of a coupling and interchange argument they provide a set of three conditions that implies optimality of a priority policy with respect to the total cost criterion. The three conditions of [60] form a subset of the parameter set that we derive for optimality of an index policy.

The paper of Down et al. [27] considers a 2-competing queues system, where the two classes have equal service rates. A coupling argument is employed to show that if two conditions hold, then an index policy is optimal for a reward variant of the model.

The approach that we will carry out is the following. First, we model the problem as a continuous time Markov decision process. To make the MDP uniformisable a truncation is necessary, after uniformisation the truncated processes can be analysed by value iteration. To show that the results for the truncated processes converge to the original model a limit theorem is needed. This theorem is only available for the discounted cost criterion, see Chapter 3. Therefore we will first show that the properties hold for the discounted cost criterion. Then via the vanishing discount approach the results are transferred to the average cost criterion (see Chapter 2).

As mentioned before, Down et al. [27] use a similar approach for the 2-competing queues system with equal service rates. Convergence of the truncated models to the original one relies on specific properties of the model and the special truncation, therefore they do not need to consider the discounted criterion. Due to the involved nature of their truncation, it seems unlikely that [27] can be extended to more dimensions or to heterogeneous service rates. Down et al. find two conditions implying optimality of an index policy. The results of the present chapter can be viewed as an extension of [27].

The optimality equation (CDOE) is a powerful tool for deriving optimal policies under the discounted cost criterion. If the value function has certain structural properties the CDOE implies optimality of an index policy. However, under naive truncations the structural properties are destroyed due to boundary effects. In [27] this problem has been solved by incorporating a specific truncation. In this chapter we use the truncation technique, called smoothed rate truncation (SRT). This technique has been introduced by Bhulai et al. [11] and can be utilised to make a process uniformisable, while keeping the structural properties intact. This method allow to show optimality of the $c\mu/\beta$-rule for the smoothed rate truncated problem. Applying the theory from Chapter 3 yields that the same policy is optimal for the unbounded problem. The theory from Chapter 2 can then be invoked to deduce that the $c\mu/\beta$-rule is optimal for the average cost criterion.

The chapter is organised as follows. In Section 5.2 we give a complete description of the model and we present the main results: if the parameters satisfy three easy conditions, then an index policy is optimal. We then identify the structural properties of the value function that imply optimality of the index policy. Section 5.3 contains the core of our analysis. First, it describes the smoothed rate truncation in more detail. Then, the structural properties of the value function are derived. In Section 5.4 we prove the main theorem. This can be done by invoking the limit theorems of Chapters 2 and 3. Section 5.5 presents some numerical examples that show that none of the three conditions is redundant. In Section 5.6 we provide the proofs of the propositions in Section 5.3.

## 5.2 Modelling and main result

### 5.2.1 Problem formulation

We consider $K$ stations that are served by a single server. We will refer to customers in station $i$ as class $i$ customers. Customers arrive to the stations according to independent Poisson processes with rates $\lambda_i > 0$ for $i = 1, \ldots, K$, respectively. The service requirement of class $i$ customers is exponentially distributed with parameter $\mu_i > 0$, $i = 1, \ldots, K$. Customers have limited patience: they are willing to wait an exponential time with parameter $\beta_i > 0$ for class $i$. We allow abandonment during service as well, resulting in an abandonment rate in station $i$ of size $\beta_i x_i$ if there are $x_i$ customers present at station $i$. Service requirements and customer impatience are mutually independent across different classes and customers.

Class $i$ customers carry holding costs $c_i \geq 0$ per unit time, $i = 1, \ldots, K$. The service regime is pre-emptive. The goal is to find the policy that minimises the average expected cost. See Section 5.2.2 for a discussion on alternative modelling choices.

We will solve this problem in the framework of Markov decision theory. To this end let the state space be $\mathbf{S} = \mathbb{N}_0^K$. Let the action space be $\mathcal{A} = \{1, \ldots, K\}$, where action $a \in \{1, \ldots, K\}$ corresponds to assigning the server to station $i$ if $a = i$. This means we only allow idling if one or more queues are empty.

We are looking for a policy that minimises the expected average cost. It can be shown that there always exists an optimal policy within the class of stationary deterministic policies. Let $\pi \in \Pi = \{\pi : \mathbf{S} \to \mathcal{A}\}$ be a stationary deterministic policy, then the rate matrix $Q(\pi)$ and cost rate $c(\pi)$ are given by

$$q_{xy}(\pi) = \begin{cases} \lambda_i & \text{if } y = x + e_i,\ i = 1, \ldots, K, \\ x_i\beta_i + \mu_i \mathbf{1}_{\{\pi(x)=i\}} & \text{if } y = x - e_i,\ x_i > 0, i = 1, \ldots, K, \\ -\sum_{z \neq x} q_{xz}(\pi) & \text{if } y = x, \end{cases}$$

$$c_x(\pi) = \sum_i c_i x_i.$$

The problem of interest is to determine the policy $\pi \in \Pi$ that minimises the expected average cost $g(\pi)$, if it exists. The quantity $g(\pi)$ is defined by

$$g(\pi) = \limsup_{T \to \infty} \frac{1}{T} \int_{t=0}^{T} \mathbb{E}^\pi \big[ c_{X_t} \big] dt.$$

Here $X_t$ denotes the number of customers in the system at time $t$, and $\mathbb{E}^\pi$ the expectation operator, if policy $\pi$ is used.

It is not to be expected that an easy expression exists that gives a full description of the optimal policy. Therefore, in this chapter we will restrict to deduce sufficient conditions for optimality of an index policy.

## 5.2.2 Main result

The main result of this chapter is Theorem 5.2.1, it provides sufficient conditions for optimality of the smallest index policy.

**Definition 5.2.1.** The smallest index policy assigns the server to the non-empty station with the smallest index. The policy only idles if no customers are present.

**Theorem 5.2.1.** *Suppose the stations can be ordered such that, for* $1 \leq i < K$, *the following three conditions hold*

$$\begin{array}{lcc} 1. & c_i \geq c_{i+1} & (c \searrow) \\ 2. & c_i\mu_i \geq c_{i+1}\mu_{i+1} & (c\mu \searrow) \\ 3. & c_i\mu_i/\beta_i \geq c_{i+1}\mu_{i+1}/\beta_{i+1} & (c\mu/\beta \searrow). \end{array} \tag{5.1}$$

*Then the smallest index policy is average optimal.*

The proof is postponed until Section 5.4. In Section 5.5 we give examples showing that if any of the three conditions of Eq. (5.1) is omitted, the smallest index policy can fail to be optimal.

The derivation of Theorem 5.2.1 naturally gives the following extra result for $\alpha$-discount optimal policies. We say that a policy is $\alpha$-discount optimal, if it is optimal with respect to the total expected discounted cost criterion with discount factor $\alpha > 0$.

**Theorem 5.2.2.** *Let discount factor* $\alpha > 0$. *Suppose the stations are ordered such that, for* $1 \leq i \leq j \leq K$, *the three conditions of Eq. (5.1) hold. Then, the smallest index policy is* $\alpha$*-discount optimal.*

**Alternative modelling choices.** In our model the cost function is a holding cost $\sum c_i x_i$ per unit time when the system is in state $x$. In many applications a penalty (say $P_i$ for class $i$) is charged if a customer abandons the system due to impatience. Then the cost per unit time is given by $\sum_i P_i\beta_i x_i$. Substitution of $c_i = P_i\beta_i$, $i = 1, \ldots, K$ implies equivalence of these cost structures.

We have modelled the system such, that customers can leave the system while being in service. In some cases it is more realistic that they are not allowed to abandon after service has started. However, if the abandonment rates are smaller than the service rates, i.e. $\beta_i < \mu_i$ for all $i$, then our analysis is still valid after an appropriate parameter change. That is, we consider the system with service rates $\hat{\mu}_i = \mu_i - \beta_i > 0$. An abandonment or service completion of the customer in service in the revised model corresponds to a service completion in the original one.

If, for one or more classes, the abandonment rates are greater than or equal to the associated service rates, then this substitution is clearly not possible. However, serving that customer class delays the process of emptying the system. Thus, it follows directly that in this case it is never optimal to serve these classes of customers. Hence, when there are only customers of that type present, then the server should idle in order to minimise the expected average cost. Therefore, the optimal policy never serves class $i$ if $\mu_i \leq \beta_i$. For the remaining customer classes with $\mu_i > \beta_i$, the smallest index policy is optimal whenever these classes satisfy the ordering $c \searrow,\ c\hat{\mu} \searrow,\ c\hat{\mu}/\beta \searrow$.

Finally, it is possible to allow idling at all times. In this case, it can easily be shown that it cannot be optimal to have unforced idling. Therefore we ignore this option for the sake of notational convenience.

### 5.2.3 Structural properties

Let $v_\pi^\alpha$ be the discounted value function under policy $\pi$, so that $v_\pi^\alpha(x)$ represents the total expected discounted cost starting at $x \in \mathbf{S}$. I.e. for $x \in \mathbf{S}$

$$v_\pi^\alpha(x) = \int_{t=0}^{\infty} e^{-\alpha t} \mathbb{E}_x^\pi \left[ c_{X_t} \right] dt.$$

Further, let $v^\alpha$ be the optimal discounted value function, defined as

$$v^\alpha(x) = \min_{\pi \in \Pi} v_\pi^\alpha(x).$$

We refer to $v^\alpha$, simply as value function hereafter. Crucial in establishing optimality of the smallest index policy are certain properties of the value function. If $v^\alpha$ is non-decreasing ($\mathcal{I}$) and weighted Upstream Increasing ($w\mathcal{UI}$), then the smallest index policy is optimal by virtue of the discounted cost optimality equation. Therefore, we introduce the following structural properties.

**Definition 5.2.2.** A function $f : \mathbf{S} \to \mathbb{R}$ is called weighted Upstream Increasing ($w\mathcal{UI}$) if $f \in w\mathcal{UI}$, with $w\mathcal{UI}$ defined as

$$\begin{aligned} w\mathcal{UI} = \{ f : \mathbf{S} \to \mathbb{R} \mid & \mu_i (f(x + e_i + e_{i+1}) - f(x + e_{i+1})) \\ & - \mu_{i+1} (f(x + e_i + e_{i+1}) - f(x + e_i)) \geq 0,\ \text{for all } x \in \mathbf{S},\ 1 \leq i < K \}. \end{aligned}$$

A function $f : \mathbf{S} \to \mathbb{R}$ is called non-decreasing ($\mathcal{I}$) if $f \in \mathcal{I}$, with $\mathcal{I}$ defined as

$$\mathcal{I} = \{ f : \mathbf{S} \to \mathbb{R} \mid f(x + e_i) - f(x) \geq 0,\ \text{for all } x \in \mathbf{S},\ 1 \leq i \leq K \}.$$

The following lemma makes the connection between the structural properties and optimality of the smallest index policy.

**Lemma 5.2.3.** *Let the discount factor $\alpha > 0$. Suppose $v^\alpha \in w\mathcal{UI} \cap \mathcal{I}$, then the smallest index policy is $\alpha$-discount optimal.*

*Proof.* Let $v^\alpha \in w\mathcal{UI} \cap \mathcal{I}$, let $1 \leq j_1 \leq j_2 \leq K$. Suppose $x$ is such that $x_{j_1}, x_{j_2} > 0$, then $v^\alpha \in w\mathcal{UI}$ implies

$$\mu_{j_1}[v^\alpha(x - e_{j_1}) - v^\alpha(x)] \leq \mu_{j_2}[v^\alpha(x - e_{j_2}) - v^\alpha(x)].$$

Now it is straightforward to check that this model satisfies the conditions of Theorem 3.4.2. For example, a drift function of the type $V(x) = e^{\epsilon \sum_i x_i}$ works well. This implies that the discount optimal policy attains the minimum in the Continuous Discount Optimality Equation (CDOE), that is

$$\begin{aligned}\alpha v^\alpha(x) = &\sum_{i=1}^K c_i x_i + \sum_{i=1}^K \lambda_i v^\alpha(x + e_i) + \sum_{i=1}^K x_i \beta_i v^\alpha(x - e_i) \\ &+ \min_{1 \leq j \leq K} \{\mu_j [v^\alpha((x - e_j)^+) - v^\alpha(x)]\} - \sum_{i=1}^K (\lambda_i + x_i \beta_i) v^\alpha(x).\end{aligned} \tag{5.2}$$

The CDOE yields that if class $j_1$ and $j_2$ customers are both present, then class $j_1$ deserves full priority over class $j_2$.

Further, since $v^\alpha$ is non-decreasing we have for $1 \leq j \leq K$, and $x$ with $x_j > 0$ that

$$\mu_j v^\alpha(x - e_j) - \mu_j v^\alpha(x) \leq 0,$$

with 0 corresponding to the cost if an empty queue is served. Hence idling is never optimal; it is optimal to serve a customer whenever possible. We conclude that the smallest index policy is optimal. □

# 5.3 Discrete time discounted cost analysis

## 5.3.1 Smoothed rate truncation

The abandonment rates increase linearly in the number of waiting customers, hence the transition rates are unbounded as a function of state. Thus, the system is not uniformisable. To make discrete time theory available, we approximate the Markov decision process with a sequence of (essentially) finite state state MDPs. Unfortunately, in standard state space truncations the structural properties of interest are lost; due to boundary effects the truncated MDP does not posses the desired structure.

To this end, the smoothed rate truncation (SRT) has been introduced in Bhulai et al. [11]. In that paper, smoothed rate truncation is applied to a Markov cost process and properties of the value function are proven.

The distinguishing feature of smoothed rate truncation is that the transition rates are decreased in all states, also close to the origin. This makes the jump rates highly state dependent and complicates the analysis, but it is the key feature of SRT that ensures that the desired properties are preserved.

The idea of SRT is as follows. Every transition that moves the system into a higher state in one or more dimensions is linearly decreased as a function of these coordinates. As we get closer to the boundary of the finite set, the rates are smoothly truncated to 0. In this way the state space is naturally restricted to a finite set of essential states, i.e. a finite number of recurrent states. On the finite state space, the transition rates are bounded. Outside the finite set the rates can be arbitrary chosen, since these states are inessential. In particular they can be chosen such that the jump rates are uniformly bounded.

In our model the truncation parameter $N = (N_1, \ldots, N_K) \in \mathcal{N} = (\mathbb{N} \cup \infty)^K$ defines the size of the state space. The set of essential states is given by $\mathbf{S}^N = \{x \in \mathbf{S} | x_i \leq N_i, i = 1, \ldots, K\}$. SRT prescribes a truncation of all transitions that move the system into a larger state. In this model only arrivals move the system to a higher state, hence the arrival rates $\lambda_i$ are replaced by new rates $\lambda_i^N(x)$ in state $x$, for all $i$. For $N_i < \infty$, the smoothed arrival rates are given by

$$\lambda_i^N(x) := (1 - \frac{x_i}{N_i})^+ \lambda_i.$$

The result is a uniformisable MDP for each finite $N \in \mathcal{N}$. Let $N = \infty^K$ correspond to the original model. This leads to a collection of parametrised Markov decision processes $\{X^N\}_{N \in \mathcal{N}}$. For $\pi \in \Pi$, $N \in \mathbb{N}^K$, the transition rates are given by

$$q_{xy}^N(\pi) = \begin{cases} \lambda_i^N(x) & \text{if } y = x + e_i,\ i = 1, \ldots, K, \\ \min\{x_i, N_i\}\beta_i + \mu_i \mathbf{1}_{\{\pi(x) = i\}} & \text{if } y = x - e_i,\ x_i > 0,\ i = 1, \ldots, K, \\ -\sum_{x \neq z} q_{xz}^N(\pi) & \text{else.} \end{cases}$$

Notice that outside $\mathbf{S}^N$ it is possible to choose the rates as we like, for these state are inessential. In particular, we can choose the new abandonment rates of class $i$ to be bounded by $N_i \beta_i$. However, it is not necessary to consider these states for determining the structural properties on $\mathbf{S}^N$.

### 5.3.2 Dynamic programming

Consider the following subset of the parameter space. Let $\mathcal{N}(\lambda)$ be given by

$$\mathcal{N}(\lambda) = \Big\{N \in \mathcal{N} \Big| \frac{\lambda_i}{N_i} \le \frac{\lambda_{i+1}}{N_{i+1}} \text{ for } 1 \le i < K\Big\}.$$

Throughout the rest of this section we fix the truncation parameter $N \in \mathcal{N}$ and discount factor $\alpha > 0$. Let $v^{\alpha,N}$ be $N$-truncated $\alpha$-discounted value function. Our goal is to show that $v^{\alpha,N} \in w\mathcal{UI} \cap \mathcal{I}$ for all $\alpha > 0$ and $N \in \mathcal{N}(\lambda)$. Use the short-hand notation for the transition rates

$$\bar{\lambda} := \sum_{i=1}^{K} \lambda_i, \qquad \beta_N := \sum_{i=1}^{K} \beta_i N_i, \qquad \mu := \max_{1\le i\le K} \mu_i.$$

Without loss of generality we assume $\bar{\lambda} + \beta_N + \mu = 1$. Then we can apply uniformisation. The discrete time MDP is then defined by

$$P^N(\pi) = I + Q^N(\pi), \quad \bar{c}(\pi) = \frac{c(\pi)}{\alpha+1}, \quad \bar{\alpha} = \frac{\alpha}{\alpha+1}, \qquad \pi \in \Pi.$$

Let $\bar{v}^{\bar{\alpha},N}$ denote the expected discrete time discount cost under the discrete time discount factor $0 < \bar{\alpha} < 1$. Then $\bar{v}^{\bar{\alpha},N} = v^{\alpha,N}$. We can approximate $\bar{v}^{\bar{\alpha},N}$ by the value iteration algorithm. Let $v_n^{\bar{\alpha},N} : \mathbf{S} \to \mathbb{R}$ for $n \ge 0$ be given by the following iteration scheme. Put $v_0^{\bar{\alpha},N} \equiv 0$, and

$$\begin{aligned}
v_{n+1}^{\bar{\alpha},N}(x) =& \sum_{i=1}^{K} \bar{c}_i x_i \quad + (1-\bar{\alpha})\Big\{ \sum_{i=1}^{K} (1 - \frac{x_i}{N_i})^+ \lambda_i v_n^{\bar{\alpha},N}(x+e_i) \\
&+ \sum_{i=1}^{K} \min\{x_i, N_i\} \beta_i v_n^{\bar{\alpha},N}(x-e_i) \\
&+ \min_{0\le j\le K} \{\mu_j [v_n^{\bar{\alpha},N}((x-e_j)^+) - v_n^{\bar{\alpha},N}(x)]\} \\
&+ \big( \sum_{i=1}^{K} \big( \min\{\frac{x_i}{N_i}, 1\}\lambda_i + (N_i - x_i)^+ \beta_i \big) + \mu \big) v_n^{\bar{\alpha},N}(x) \Big\} \\
=&(1-\bar{\alpha})\Big( \sum_{i=1}^{K} c_i x_i + \Big\{ \sum_{i=1}^{K} (1 - \frac{x_i}{N_i})^+ \lambda_i v_n^{\bar{\alpha},N}(x+e_i) \\
&+ \sum_{i=1}^{K} \min\{x_i, N_i\} \beta_i v_n^{\bar{\alpha},N}(x-e_i)
\end{aligned}$$

$$+ \min_{0\le j\le K} \{\mu_j [v_n^{\bar{\alpha},N}((x-e_j)^+) - v_n^{\bar{\alpha},N}(x)]\}$$
$$+ \Big(\sum_{i=1}^{K} \big(\min\{\frac{x_i}{N_i},1\}\lambda_i + (N_i - x_i)^+\beta_i\big) + \mu\Big) v_n^{\bar{\alpha},N}(x)\Big\}\Big).$$

We will prove by induction that $v_n^{\bar{\alpha},N} \in w\mathcal{UI} \cap \mathcal{I}$ on $\mathbf{S}^N$, for all $n \ge 0$. To employ the induction argument we need three additional structural properties: convexity, supermodularity and bounded increasingness. We will specify these hereafter. The induction hypothesis $v_0^{\bar{\alpha},N} \equiv 0$ trivially satisfies all these properties. For the induction step we are not going to look at the iteration in its entirety, but we will use event based dynamic programming (EBDP). This method uses event operators – representing arrivals, departures or costs – as building blocks to construct the iteration step of the value iteration algorithm.

**Definition 5.3.1.** Let $f : \mathbf{S} \to \mathbb{R}$, then define

1. a) The total smoothed arrivals operator

$$\mathcal{T}_{SA}^N f := \bar{\lambda}^{-1} \sum_{i=1}^{K} \lambda_i \mathcal{T}_{SA(i)}^N f,$$

   b) with the smoothed arrivals operator given by

$$\mathcal{T}_{SA(i)}^N f(x) := \begin{cases} (1-\frac{x_i}{N_i}) f(x+e_i) + \frac{x_i}{N_i} f(x), & x_i \le N_i, \\ f(x), & \text{else.} \end{cases}$$

2. a) The total increasing departures operator

$$\mathcal{T}_{ID}^N f := \beta_N^{-1} \sum_{i=1}^{K} \beta_i N_i \mathcal{T}_{ID(i)}^N f,$$

   b) with the increasing departures operator

$$\mathcal{T}_{ID(i)}^N f(x) := \begin{cases} \frac{x_i}{N_i} f(x-e_i) + (1-\frac{x_i}{N_i}) f(x), & x_i \le N_i, \\ f(x-e_i), & \text{else.} \end{cases}$$

3. The cost operator

$$\mathcal{T}_C f(x) := \sum_{i=1}^{K} c_i x_i + f(x).$$

4. The cost + increasing departures operator

$$\mathcal{T}_{CID}^{N} := \beta_N^{-1}\mathcal{T}_C(\beta_N \mathcal{T}_{ID}^{N}).$$

5. The movable server operator

$$\mathcal{T}_{MS} f(x) := \min_{1 \le j \le K} \Big\{ \frac{\mu_j}{\mu} f((x-e_j)^+) + (1 - \frac{\mu_j}{\mu}) f(x) \Big\}.$$

6. For $f_1, f_2, f_3 : \mathbf{S} \to \mathbb{R}$, the uniformisation operator

$$\mathcal{T}_{UNIF}(f_1, f_2, f_3) := \bar{\lambda} f_1 + \beta_N f_2 + \mu f_3.$$

7. The discount operator

$$\mathcal{T}_{DISC}^{\bar{\alpha}} f := (1-\bar{\alpha}) f.$$

Now $v_{n+1}^{\bar{\alpha},N}$ can be constructed as follows

$$\mathcal{T}_{DISC}^{\bar{\alpha}}(\mathcal{T}_{UNIF}(\mathcal{T}_{SA}^{N} v_n^{\bar{\alpha},N}, \mathcal{T}_{CID}^{N} v_n^{\bar{\alpha},N}, \mathcal{T}_{MS} v_n^{\bar{\alpha},N})) = v_{n+1}^{\bar{\alpha},N}.$$

It is sufficient that $v_n^{\bar{\alpha},N}$ has the desired structural properties on the essential states, the finite set $\mathbf{S}^N$. Therefore define the following collections of functions that possess a certain property restricted to $\mathbf{S}^N$.

**Definition 5.3.2.** (Properties on $\mathbf{S}^N$)

1. Weighted upstream increasing functions on $\mathbf{S}^N$

$$\begin{aligned} w\mathcal{UI}^N = \{ f : \mathbf{S} \to \mathbb{R} \mid & \mu_i (f(x+e_i+e_{i+1}) - f(x+e_{i+1})) \\ & -\mu_{i+1}(f(x+e_i+e_{i+1}) - f(x+e_i)) \ge 0, \\ & \text{for all } x, x+e_i+e_{i+1} \in \mathbf{S}^N,\ 1 \le i < K \}. \end{aligned}$$

2. Increasing functions on $\mathbf{S}^N$

$$\begin{aligned} \mathcal{I}^N = \{ f : \mathbf{S} \to \mathbb{R} \mid & f(x+e_i) - f(x) \ge 0, \\ & \text{for all } x, x+e_i \in \mathbf{S}^N,\ 1 \le i \le K \}. \end{aligned}$$

3. Supermodular functions on $\mathbf{S}^N$

$$\begin{aligned} \mathcal{S}uper^N = \{ f : \mathbf{S} \to \mathbb{R} \mid & f(x+e_i+e_j) - f(x+e_i) \\ & -f(x+e_j) + f(x) \ge 0, \\ & \text{for all } x, x+e_i+e_j \in \mathbf{S}^N,\ 1 \le i \ne j \le K \}. \end{aligned}$$

4. Convex functions on $\mathbf{S}^N$

$$\mathcal{C}x^N = \{f : \mathbf{S} \to \mathbb{R} \mid f(x+2e_i) - 2f(x+e_i) + f(x) \geq 0, \\ \text{for all } x, x+2e_i \in \mathbf{S}^N,\ 1 \leq i \leq K\}.$$

5. Bounded increasing functions on $\mathbf{S}^N$

$$\mathcal{BD}^N = \{f : \mathbf{S} \to \mathbb{R} \mid f(x+e_i) - f(x) \leq \frac{c_i}{\beta_i}, \\ \text{for all } x, x+e_i \in \mathbf{S}^N.\ 1 \leq i \leq K\}.$$

The following propositions are sufficient for the induction step.

**Proposition 5.3.1.** *The smoothed arrivals operator has the following propagation properties*

*i)*

$$\mathcal{T}_{SA}^N : \mathcal{I}^N \to \mathcal{I}^N, \mathcal{C}x^N \to \mathcal{C}x^N, \mathcal{S}uper^N \to \mathcal{S}uper^N, \mathcal{BD}^N \to \mathcal{BD}^N.$$

*ii) If moreover* $N \in \mathcal{N}(\lambda)$, *then*

$$\mathcal{T}_{SA}^N : \mathcal{I}^N \cap w\mathcal{UI}^N \to w\mathcal{UI}^N.$$

**Proposition 5.3.2.** *The increasing departure operator has the following propagation properties*

$$\mathcal{T}_{ID}^N : \mathcal{I}^N \to \mathcal{I}^N, \mathcal{C}x^N \to \mathcal{C}x^N, \mathcal{S}uper^N \to \mathcal{S}uper^N.$$

**Proposition 5.3.3.** *The cost operator has the following propagation properties*

$$\mathcal{T}_C : \mathcal{I}^N \to \mathcal{I}^N, \mathcal{C}x^N \to \mathcal{C}x^N, \mathcal{S}uper^N \to \mathcal{S}uper^N.$$

**Proposition 5.3.4.** *The cost + increasing departures operator has the following propagation properties*

*i)*

$$\mathcal{T}_{CID}^N : \mathcal{BD}^N \to \mathcal{BD}^N.$$

*ii) If moreover, for all* $1 \leq i < K$, *Eq. (5.1) holds, then*

$$\mathcal{T}_{CID}^N : \mathcal{I}^N \cap w\mathcal{UI}^N \cap \mathcal{S}uper^N \cap \mathcal{BD}^N \to w\mathcal{UI}^N.$$

**Proposition 5.3.5.** *The movable server operator has the following propagation properties*

*i)*

$$\begin{aligned}\mathcal{T}_{MS} : \mathcal{I}^N \cap w\mathcal{UI}^N &\to \mathcal{I}^N \cap w\mathcal{UI}^N, \\ \mathcal{I}^N \cap w\mathcal{UI}^N \cap \mathcal{C}x^N \cap \mathcal{S}uper^N &\to \mathcal{C}x^N \cap \mathcal{S}uper^N.\end{aligned}$$

*ii) If moreover, for all $1 \le i < K$, $c_i\mu_i/\beta_i \ge c_{i+1}\mu_{i+1}/\beta_{i+1}$ then*

$$\mathcal{T}_{MS} : \mathcal{I}^N \cap w\mathcal{UI}^N \cap \mathcal{BD}^N \to \mathcal{BD}^N.$$

**Proposition 5.3.6.** *The uniformisation operator has the following propagation properties:*

$$\begin{aligned}\mathcal{T}_{UNIF} : (w\mathcal{UI}^N)^3 \to w\mathcal{UI}^N, (\mathcal{I}^N)^3 \to \mathcal{I}^N, (\mathcal{C}x^N)^3 \to \mathcal{C}x^N, \\ (\mathcal{S}uper^N)^3 \to \mathcal{S}uper^N, (\mathcal{BD}^N)^3 \to \mathcal{BD}^N.\end{aligned}$$

**Proposition 5.3.7.** *The discount operator has the following propagation properties:*

$$\begin{aligned}\mathcal{T}^{\bar{\alpha}}_{DISC} : w\mathcal{UI}^N \to w\mathcal{UI}^N, \mathcal{I}^N \to \mathcal{I}^N, \mathcal{C}x^N \to \mathcal{C}x^N, \\ \mathcal{S}uper^N \to \mathcal{S}uper^N, \mathcal{BD}^N \to \mathcal{BD}^N.\end{aligned}$$

The proofs of the propositions are placed in Section 5.6.

**Corollary 5.3.8.** *Let $N \in \mathcal{N}(\lambda)$, $0 < \bar{\alpha} < 1$ and suppose that for all $1 \le i < K$, Eq. (5.1) holds.*

*i) Then for all $n \ge 0$*

$$v_n^{\bar{\alpha},N} \in w\mathcal{UI}^N \cap \mathcal{I}^N \cap \mathcal{C}x^N \cap \mathcal{S}uper^N \cap \mathcal{BD}^N;$$

*ii) furthermore*

$$\bar{v}^{\bar{\alpha},N} \in w\mathcal{UI}^N \cap \mathcal{I}^N \cap \mathcal{C}x^N \cap \mathcal{S}uper^N \cap \mathcal{BD}^N.$$

*Proof.* Denote $\mathcal{A} = w\mathcal{UI}^N \cap \mathcal{I}^N \cap \mathcal{C}x^N \cap \mathcal{S}uper^N \cap \mathcal{BD}^N$. First notice that $v_0^{\bar{\alpha},N} \in \mathcal{A}$. Further, under above conditions we have $\mathcal{T}^N_{SA}$, $\mathcal{T}^N_{CID}$, $\mathcal{T}_{MS}$, $\mathcal{T}^{\bar{\alpha}}_{DISC} : \mathcal{A} \to \mathcal{A}$ and $\mathcal{T}_{UNIF} : \mathcal{A}^3 \to \mathcal{A}$. This means that

$$\mathcal{T}^{\bar{\alpha}}_{DISC}(\mathcal{T}_{UNIF}(\mathcal{T}^N_{SA}, \mathcal{T}^N_{CID}, \mathcal{T}_{MS})) : \mathcal{A} \to \mathcal{A}.$$

Now suppose that $v_n^{\bar{\alpha},N} \in \mathcal{A}$, the above implies that

$$v_{n+1}^{\bar{\alpha},N} = \mathcal{T}^{\bar{\alpha}}_{DISC}(\mathcal{T}_{UNIF}(\mathcal{T}^N_{SA}, \mathcal{T}^N_{CID}, \mathcal{T}_{MS}))v_n^{\bar{\alpha},N} \in \mathcal{A}.$$

Now, Assertion i) follows by induction.

Assertion ii) immediately follows from i) due to convergence of value iteration (see Theorem 2.2.3). □

## 5.4 Proof of main theorems

*Proof of Theorem 5.2.2.* Suppose that for all $1 \leq i < K$, Eq. (5.1) holds. Let the continuous time discount factor $\alpha > 0$, then the discrete time discount factor $\bar{\alpha} = \alpha/(\alpha+1)$ satisfies $0 < \bar{\alpha} < 1$. Take $N \in \mathcal{N}(\lambda)$, then Corollary 5.3.8 implies that $\bar{v}^{\bar{\alpha},N} = v^{\alpha,N} \in w\mathcal{UI}^N \cap \mathcal{I}^N$. This model satisfies the assumptions of Theorem 3.5.1 on parametrised Markov processes, completely analogous to the applicability of CDOE (5.2) by means of Theorem 3.4.2. Theorem 3.5.1 implies continuity in the truncation parameter. This means that $v^{\alpha,N} \to v^\alpha$ as $N_i \to \infty$, for $i = 1, \ldots, K$. Hence, $v^\alpha \in w\mathcal{UI} \cap \mathcal{I}$ and therefore by Lemma 5.2.3 the smallest index policy is $\alpha$-discount optimal. $\square$

*Proof of Theorem 5.2.1.* Suppose that for all $1 \leq i < K$, Eq. (5.1) holds. By Theorem 5.2.2 we have that for all $\alpha > 0$ that the smallest index policy is $\alpha$-discount optimal, denote this policy as $\pi^\alpha$.

Notice that the model satisfies the assumptions of Theorem 2.2.5. This is straightforward because 0 is an ergodic state for every policy and for all $N \in \mathcal{N}$. This theorem implies the existence of a sequence $(\alpha_m)$ with $\lim_{m\to\infty} \alpha_m = 0$, such that the limit $\lim_{m\to\infty} \pi^{\alpha_m}$ is average optimal. Since $\pi^\alpha$ is the smallest index policy for all $\alpha$, so is the limit policy. Hence the smallest index policy is average optimal. $\square$

## 5.5 Numerical results

The triple set of inequalities, under which the smallest index policy is optimal, induces a lot of parameter configurations that fall outside the scope of the theorems. This naturally gives rise to the question whether all three conditions are necessary.

From numerical calculations, it follows that we cannot omit one of the three conditions. If one of these three inequalities is violated then the examples below show that the smallest index policy need not be optimal. We carried out the calculations for $K = 2$. In order to keep the structures intact we have used smoothed rate truncation.

1. Consider the following parameter setting:

| $c\uparrow$ | $c_i$ | $c_i\mu_i$ | $c_i\mu_i/\beta_i$ | $\lambda_i$ | $\mu_i$ | $\beta_i$ |
|---|---|---|---|---|---|---|
| $i=1$ | 0.025 | 1.25 | 12.02 | 2 | 50 | 0.104 |
| $i=2$ | 1.2 | 1.2 | 12 | 2 | 1 | 0.1 |

We see that Eq. (5.1)-1 is violated, $c_1 < c_2$, while Eqs. (5.1)-2 and (5.1)-3 are satisfied. The optimal policy is a switching curve policy: for small states action 1 is optimal and for large states action 2 is optimal, see Table 5.1.

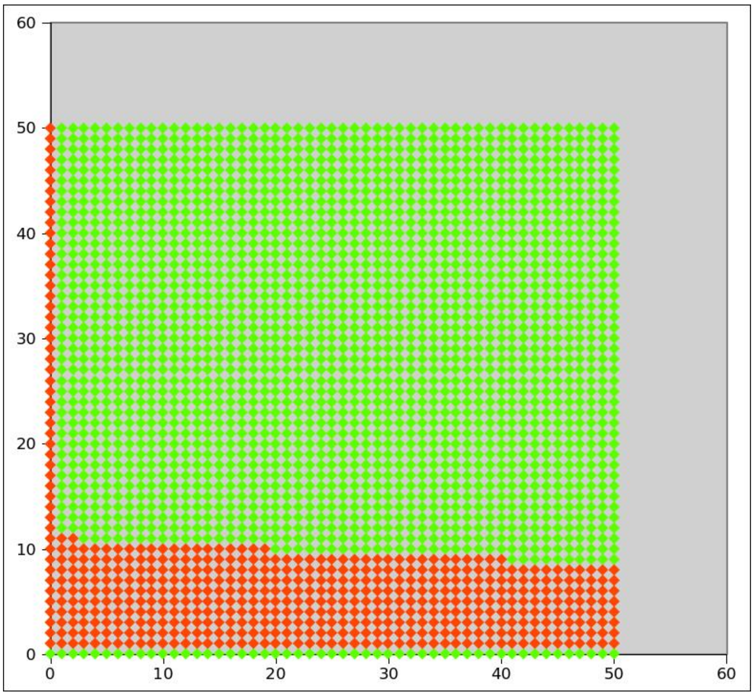



Table 5.1: Optimality of a switching curve if Eq. (5.1)-1 is violated

2. The next parameter setting is given by:

| $c\mu \uparrow$ | $c_i$ | $c_i\mu_i$ | $c_i\mu_i/\beta_i$ | $\lambda_i$ | $\mu_i$ | $\beta_i$ |
|---|---|---|---|---|---|---|
| $i=1$ | 1 | 1 | 100 | 0.5 | 1 | 0.01 |
| $i=2$ | 1 | 2 | 20 | 0.5 | 2 | 0.01 |

Observe that Eqs. (5.1)-1 and (5.1)-3 hold, but Eq. (5.1)-2 is violated. Table 5.2 displays the optimal policy. We see that the smallest index policy need not be optimal. There is only a small region – if there are only few customers in the system– where it is optimal to serve the customer with the smallest index. In larger states action 2 is optimal.

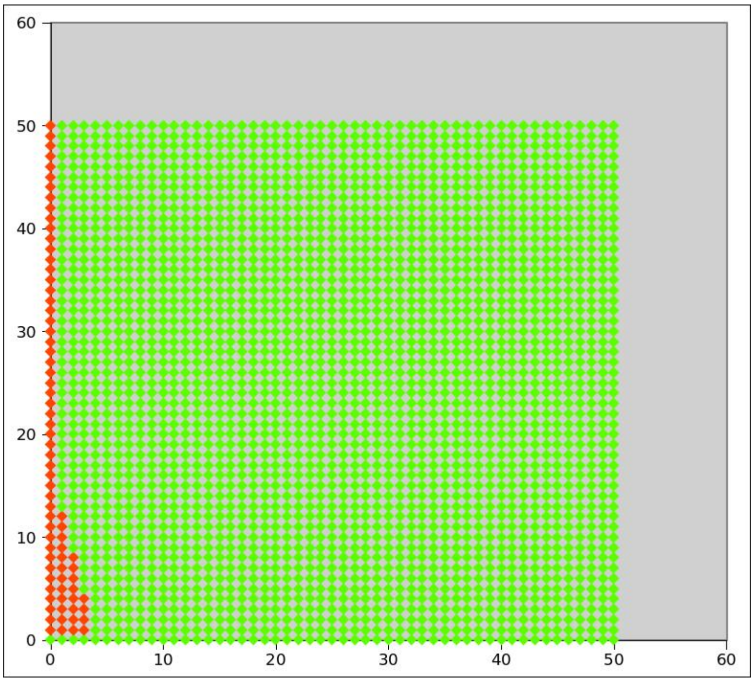



Table 5.2: Optimality of a switching curve if Eq. (5.1)-2 is violated

3. The third parameter setting is:

| $c\mu/\beta\uparrow$ | $c_i$ | $c_i\mu_i$ | $c_i\mu_i/\beta_i$ | $\lambda_i$ | $\mu_i$ | $\beta_i$ |
|---|---|---|---|---|---|---|
| $i=1$ | 1.2 | 1.2 | 2.4 | 2 | 1 | 0.5 |
| $i=2$ | 1 | 1 | 2.5 | 2 | 1 | 0.4 |

Here only Eqs. (5.1)-1 and (5.1)-2 are satisfied, Table 5.3 shows that it can be optimal to serve the station with the highest index instead of the smallest index.

Another observation can be made when considering these examples. In all cases a switching curve policy is optimal, if we view an index policy as a degenerate switching curve policy. We conjecture that a switching curve policy is always optimal.

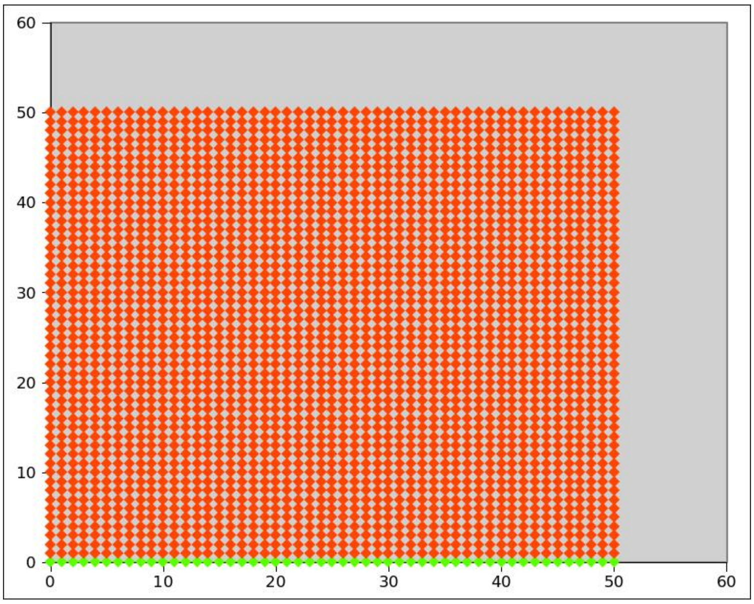



Table 5.3: Optimality of a highest index policy if Eq. (5.1)-3 is violated

## 5.6 Proofs of propagation results

In this section we will provide the proofs of Propositions 5.3.1-5.3.7. We make use of the following notation.

**Definition 5.6.1.** 1. For $1 \leq i < K$,

$$w\mathcal{UI}^N(i) = \{f : \mathbf{S} \to \mathbb{R} \mid \mu_i(f(x + e_i + e_{i+1}) - f(x + e_{i+1})) - \mu_{i+1}(f(x + e_i + e_{i+1}) - f(x + e_i)) \geq 0, \text{ for all } x, x + e_i + e_{i+1} \in \mathbf{S}^N\}.$$

2. For $1 \leq i \leq K$,

$$\mathcal{I}^N(i) = \{f : \mathbf{S} \to \mathbb{R} \mid f(x + e_i) - f(x) \geq 0, \text{ for all } x, x + e_i \in \mathbf{S}^N\}.$$

3. For $1 \leq i \leq K$,

$$\mathcal{C}x^N(i) = \{f : \mathbf{S} \to \mathbb{R} \mid f(x + 2e_i) - 2f(x + e_i) + f(x) \geq 0, \text{ for all } x, x + 2e_i \in \mathbf{S}^N\}.$$

4. For $1 \leq i \neq j \leq K$,

$$\mathcal{S}uper^N(i,j) = \{f : \mathbf{S} \to \mathbb{R} \mid f(x + e_i + e_j) - f(x + e_i) - f(x + e_j) + f(x) \geq 0, \text{ for all } x, x + e_i + e_j \in \mathbf{S}^N\}.$$

5. For $1 \le i \le K$,

$$\mathcal{BD}^N(i) = \{f : \mathbf{S} \to \mathbb{R} \mid f(x+e_i) - f(x) \le \frac{c_i}{\beta_i}, \text{ for all } x, x+e_i \in \mathbf{S}^N\}.$$

It is straightforward, that

$$w\mathcal{UI}^N = \bigcap_{1\le i<K} w\mathcal{UI}^N(i),\ \mathcal{I}^N = \bigcap_{1\le i\le K} \mathcal{I}^N(i),\ \mathcal{C}x^N = \bigcap_{1\le i\le K} \mathcal{C}x^N(i),$$
$$\mathcal{S}uper^N = \bigcap_{1\le i\ne j\le K} \mathcal{S}uper^N(i,j),\ \mathcal{BD}^N = \bigcap_{1\le i\le K} \mathcal{BD}^N(i).$$

*Proof of Proposition 5.3.1.* First we show that $\mathcal{T}_{SA}^N : \mathcal{I}^N \to \mathcal{I}^N$. Suppose $f \in \mathcal{I}^N$, let $1 \le i \le K$, then for all $x, x+e_i \in \mathbf{S}^N$, we have

$$\begin{aligned}
&\mathcal{T}_{SA(i)}^N f(x+e_i) - \mathcal{T}_{SA(i)}^N f(x)\\
=\ & (1-\frac{x_i+1}{N_i})f(x+2e_i) + \frac{x_i+1}{N_i}f(x+e_i)\\
& -\big((1-\frac{x_i}{N_i})f(x+e_i) + \frac{x_i}{N_i}f(x)\big)\\
=\ & (1-\frac{x_i+1}{N_i})[f(x+2e_i) - f(x+e_i)] - \frac{1}{N_i}f(x+e_i)\\
& +\frac{x_i}{N_i}[f(x+e_i) - f(x)] + \frac{1}{N_i}f(x+e_i)\\
\ge\ & 0.
\end{aligned}$$

The inequality is due to $f \in \mathcal{I}^N(i)$. Hence, we obtain $\mathcal{T}_{SA(i)}^N f \in \mathcal{I}^N(i)$. Notice that $\mathcal{T}_{SA(i)}^N f \in \mathcal{I}^N(j)$ for $i \ne j$ is trivial. Hence, we conclude that $\mathcal{T}_{SA(i)}^N : \mathcal{I}^N \to \mathcal{I}^N$. This yields $\mathcal{T}_{SA}^N : \mathcal{I}^N \to \mathcal{I}^N$.

Next we prove $\mathcal{T}_{SA}^N : \mathcal{C}x^N \to \mathcal{C}x^N$. Assume that $f \in \mathcal{C}x^N$. Let $1 \le i \le K$, then $x, x+2e_i \in \mathbf{S}^N$ implies

$$\begin{aligned}
&\mathcal{T}_{SA(i)}^N f(x+2e_i) - 2\mathcal{T}_{SA(i)}^N f(x+e_i) + \mathcal{T}_{SA(i)}^N f(x)\\
=\ & (1-\frac{x_i+2}{N_i})f(x+3e_i) + \frac{x_i+2}{N_i}f(x+2e_i)\\
& -2(1-\frac{x_i+1}{N_i})f(x+2e_i) - 2\frac{x_i+1}{N_i}f(x+e_i)\\
& +(1-\frac{x_i}{N_i})f(x+e_i) + \frac{x_i}{N_i}f(x)
\end{aligned}$$

$$
\begin{aligned}
&= (1-\frac{x_i+2}{N_i})[f(x+3e_i)-2f(x+2e_i)+f(x+e_i)] \\
&\qquad -\frac{2}{N_i}f(x+2e_i)+\frac{2}{N_i}f(x+e_i) \\
&\quad +\frac{x_i}{N_i}[f(x+2e_i)-2f(x+e_i)+f(x)] \\
&\qquad +\frac{2}{N_i}f(x+2e_i)-\frac{2}{N_i}f(x+e_i) \\
&\geq 0.
\end{aligned}
$$

The inequality follows from $f \in \mathcal{C}x^N(i)$. Hence $\mathcal{T}^N_{SA(i)}f \in \mathcal{C}x^N(i)$. Trivially, we also have for $j \neq i$ that $\mathcal{T}^N_{SA(i)}f \in \mathcal{C}x^N(j)$. So $\mathcal{T}^N_{SA(i)}f \in \mathcal{C}x^N$ for any $i$, and we may conclude that $\mathcal{T}^N_{SA} : \mathcal{C}x^N \to \mathcal{C}x^N$.

Next we prove $\mathcal{T}^N_{SA} : \mathcal{S}uper^N \to \mathcal{S}uper^N$. Suppose $f \in \mathcal{S}uper^N$. Let $1 \leq i \neq j \leq K$ be arbitrary, then for all $x$ with $x, x+e_i+e_j \in \mathbf{S}^N$ we have

$$
\begin{aligned}
&\mathcal{T}^N_{SA(i)}f(x+e_i+e_j)-\mathcal{T}^N_{SA(i)}f(x+e_i)-\mathcal{T}^N_{SA(i)}f(x+e_j)+\mathcal{T}^N_{SA(i)}f(x) \\
&= (1-\frac{x_i+1}{N_i})f(x+2e_i+e_j)+\frac{x_i+1}{N_i}f(x+e_i+e_j) \\
&\quad -(1-\frac{x_i+1}{N_i})f(x+2e_i)-\frac{x_i+1}{N_i}f(x+e_i) \\
&\quad -(1-\frac{x_i}{N_i})f(x+e_i+e_j)-\frac{x_i}{N_i}f(x+e_j) \\
&\quad +(1-\frac{x_i}{N_i})f(x+e_i)+\frac{x_i}{N_i}f(x) \\
&= (1-\frac{x_i+1}{N_i})[f(x+2e_i+e_j)-f(x+2e_i)-f(x+e_i+e_j)+f(x+e_i)] \\
&\qquad -\frac{1}{N_i}(f(x+e_i+e_j)-f(x+e_i)) \\
&\quad +\frac{x_i}{N_i}[f(x+e_i+e_j)-f(x+e_i)-f(x+e_j)+f(x)] \\
&\qquad +\frac{1}{N_i}(f(x+e_i+e_j)-f(x+e_i)) \\
&\geq 0.
\end{aligned}
$$

We use $f \in \mathcal{S}uper^N(i,j)$ for the terms between square brackets to get the inequality. Thus $\mathcal{T}^N_{SA(i)}f \in \mathcal{S}uper^N(i,j)$. It immediately follows that $\mathcal{T}^N_{SA(i)}f \in \mathcal{S}uper^N(j,k)$, if $i \neq j,k$. Hence we have $\mathcal{T}^N_{SA(i)} : \mathcal{S}uper^N \to \mathcal{S}uper^N$ for any $i$, which implies $\mathcal{T}_{SA}^N : \mathcal{S}uper^N \to \mathcal{S}uper^N$.

Next we prove $\mathcal{T}_{SA}^N : \mathcal{BD}^N \to \mathcal{BD}^N$. Suppose that $f \in \mathcal{BD}^N$. Let $1 \leq i \leq K$, then for all $x$, with $x, x+e_i \in \mathbf{S}^N$ we have

$$\begin{aligned}
&\mathcal{T}_{SA(i)}^N f(x+e_i) - \mathcal{T}_{SA(i)}^N f(x) \\
&\quad = (1 - \frac{x_i+1}{N_i}) f(x+2e_i) + \frac{x_i+1}{N_i} f(x+e_i) \\
&\qquad -(1 - \frac{x_i}{N_i}) f(x+e_i) - \frac{x_i}{N_i} f(x) \\
&\quad = (1 - \frac{x_i+1}{N_i})[f(x+2e_i) - f(x+e_i)] - \frac{1}{N_i} f(x+e_i) \\
&\qquad +\frac{x_i}{N_i}[f(x+e_i) - f(x)] + \frac{1}{N_i} f(x+e_i) \\
&\quad \leq (1 - \frac{1}{N_i}) \frac{c_i}{\beta_i} \\
&\quad \leq \frac{c_i}{\beta_i}.
\end{aligned}$$

We use $f \in \mathcal{BD}^N(i)$ for the terms in square brackets, to obtain the first inequality. Hence, $\mathcal{T}_{SA(i)}^N f \in \mathcal{BD}^N(i)$. It easily follows that $\mathcal{T}_{SA(i)}^N : \mathcal{BD}^N \to \mathcal{BD}^N$, and so $\mathcal{T}_{SA}^N : \mathcal{BD}^N \to \mathcal{BD}^N$.

For the proof of *ii*) assume $N \in \mathcal{N}(\lambda)$, so that $\lambda_{i+1}/N_{i+1} \geq \lambda_i/N_i$ for $1 \leq i < K$. We will prove that $\mathcal{T}_{SA}^N : \mathcal{I}^N \cap w\mathcal{UI}^N \to w\mathcal{UI}^N$. The property $w\mathcal{UI}^N$ does not propagate through one individual smoothed arrivals operator, and so it is necessary to look at the combined smoothed arrivals operator $\mathcal{T}_{SA}^N$. Suppose, that $f \in \mathcal{I}^N \cap w\mathcal{UI}^N$. It suffices to show, that $\mathcal{T}_{SA}^N f \in w\mathcal{UI}^N(i)$, for an arbitrary $1 \leq i < K$. First, consider the $\mathcal{T}_{SA(i)}^N$ operator for $x$ with $x, x+e_i+e_{i+1} \in \mathbf{S}^N$. Then,

$$\begin{aligned}
&\mu_i(\mathcal{T}_{SA(i)}^N f(x+e_i+e_{i+1}) - \mathcal{T}_{SA(i)}^N f(x+e_{i+1})) \\
&\qquad - \mu_{i+1}(\mathcal{T}_{SA(i)}^N f(x+e_i+e_{i+1}) - \mathcal{T}_{SA(i)}^N f(x+e_i)) \\
&= \mu_i\big((1 - \frac{x_i+1}{N_i}) f(x+2e_i+e_{i+1}) + \frac{x_i+1}{N_i} f(x+e_i+e_{i+1})\big) \\
&\qquad - \mu_i\big((1 - \frac{x_i}{N_i}) f(x+e_i+e_{i+1}) + \frac{x_i}{N_i} f(x+e_{i+1})\big) \\
&\qquad - \mu_{i+1}\big((1 - \frac{x_i+1}{N_i}) f(x+2e_i+e_{i+1}) + \frac{x_i+1}{N_i} f(x+e_i+e_{i+1})\big)
\end{aligned}$$

$$
\begin{aligned}
&\quad + \mu_{i+1}\big((1-\frac{x_i+1}{N_i})f(x+2e_i)+\frac{x_i+1}{N_i}f(x+e_i)\big)\\
= &\quad (1-\frac{x_i+1}{N_i})\big[\mu_i(f(x+2e_i+e_{i+1})-f(x+e_i+e_{i+1}))\\
&\qquad\qquad\qquad - \mu_{i+1}(f(x+2e_i+e_{i+1})-f(x+2e_i))\big]\\
&\quad + \frac{x_i}{N_i}\big[\mu_i(f(x+e_i+e_{i+1})-f(x+e_{i+1}))\\
&\qquad\qquad\qquad - \mu_{i+1}(f(x+e_i+e_{i+1})-f(x+e_i))\big]\\
&\quad - \frac{1}{N_i}\mu_{i+1}(f(x+e_i+e_{i+1})-f(x+e_i))\\
\geq &\quad - \frac{1}{N_i}\mu_{i+1}(f(x+e_i+e_{i+1})-f(x+e_i)). \qquad (5.3)
\end{aligned}
$$

The terms between square brackets are greater or equal to zero because $f \in w\mathcal{UI}^N(i)$. Notice that the resulting term is smaller or equal than zero, since $f \in \mathcal{I}^N$. Combine this with the $\mathcal{T}^N_{SA(i+1)}$ operator. Then similar as in the above we get

$$
\begin{aligned}
&\mu_i(\mathcal{T}^N_{SA(i+1)}f(x+e_i+e_{i+1})-\mathcal{T}^N_{SA(i+1)}f(x+e_{i+1}))\\
&\qquad -\mu_{i+1}(\mathcal{T}^N_{SA(i+1)}f(x+e_i+e_{i+1})-\mathcal{T}^N_{SA(i+1)}f(x+e_i))\\
&\quad = \quad (1-\frac{x_{i+1}+1}{N_{i+1}})\big[\mu_i(f(x+e_i+2e_{i+1})-f(x+2e_{i+1}))\\
&\qquad\qquad\qquad -\mu_{i+1}(f(x+e_i+2e_{i+1})-f(x+e_i+e_{i+1}))\big]\\
&\qquad\quad +\frac{x_{i+1}}{N_{i+1}}\big[\mu_i(f(x+e_i+e_{i+1})-f(x+e_{i+1}))\\
&\qquad\qquad\qquad -\mu_{i+1}(f(x+e_i+e_{i+1})-f(x+e_i))\big]\\
&\qquad\quad +\frac{1}{N_{i+1}}\mu_i(f(x+e_i+e_{i+1})-f(x+e_{i+1}))\\
&\quad \geq \quad \frac{1}{N_{i+1}}\mu_i(f(x+e_i+e_{i+1})-f(x+e_{i+1})). \qquad (5.4)
\end{aligned}
$$

Then we obtain for the total smoothed arrivals operator

$$
\begin{aligned}
&\bar{\lambda}\Big(\mu_i(\mathcal{T}^N_{SA}f(x+e_i+e_{i+1})-\mathcal{T}^N_{SA}f(x+e_{i+1}))\\
&\qquad -\mu_{i+1}(\mathcal{T}^N_{SA}f(x+e_i+e_{i+1})-\mathcal{T}^N_{SA}f(x+e_i))\Big)\\
&\quad \geq \quad \lambda_i\Big(\mu_i(\mathcal{T}^N_{SA(i)}f(x+e_i+e_{i+1})-\mathcal{T}^N_{SA(i)}f(x+e_{i+1}))
\end{aligned}
$$

$$
\begin{aligned}
& \quad -\mu_{i+1}(\mathcal{T}^N_{SA(i)}f(x+e_i+e_{i+1}) - \mathcal{T}^N_{SA(i)}f(x+e_i))\Big) \\
+ \quad & \lambda_{i+1}\Big(\mu_i(\mathcal{T}^N_{SA(i+1)}f(x+e_i+e_{i+1}) - \mathcal{T}^N_{SA(i+1)}f(x+e_{i+1})) \\
& \quad -\mu_{i+1}(\mathcal{T}^N_{SA(i+1)}f(x+e_i+e_{i+1}) - \mathcal{T}^N_{SA(i+1)}f(x+e_i))\Big) \\
\geq \quad & \frac{\lambda_{i+1}}{N_{i+1}}\mu_i(f(x+e_i+e_{i+1}) - f(x+e_{i+1})) \\
& -\frac{\lambda_i}{N_i}\mu_{i+1}(f(x+e_i+e_{i+1}) - f(x+e_i)) \\
= \quad & \frac{\lambda_i}{N_i}\big[\mu_i(f(x+e_i+e_{i+1}) - f(x+e_{i+1})) \\
& \qquad\qquad -\mu_{i+1}(f(x+e_i+e_{i+1}) - f(x+e_i))\big] \\
& +\left(\frac{\lambda_{i+1}}{N_{i+1}} - \frac{\lambda_i}{N_i}\right)\mu_i[f(x+e_i+e_{i+1}) - f(x+e_{i+1})] \\
\geq \quad & 0.
\end{aligned}
$$

The first inequality is due to the fact that $\mathcal{T}^N_{SA(j)}$ for $j \neq i, i+1$ trivially propagates $w\mathcal{UI}^N(i)$. The second inequality follows from Inequalities (5.3) and (5.4). The third inequality follows from $f \in \mathcal{I}^N \cap w\mathcal{UI}^N$ and $\frac{\lambda_{i+1}}{N_{i+1}} \geq \frac{\lambda_i}{N_i}$. So we have $\mathcal{T}^N_{SA}f \in w\mathcal{UI}^N(i)$ for every $1 \leq i < K$, hence $\mathcal{T}^N_{SA} : \mathcal{I}^N \cap w\mathcal{UI}^N \to w\mathcal{UI}^N$. □

*Proof of Proposition 5.3.2.* Start with the proof of $\mathcal{T}^N_{ID} : \mathcal{I}^N \to \mathcal{I}^N$. Suppose, that $f \in \mathcal{I}^N$, let $1 \leq i \leq K$ be arbitrary, then for all $x$ such that $x, x+e_i \in \mathbf{S}^N$ we have

$$
\begin{aligned}
& \mathcal{T}^N_{ID(i)}f(x+e_i) - \mathcal{T}^N_{ID(i)}f(x) \\
& \quad = \quad \frac{x_i+1}{N_i}f(x) + (1 - \frac{x_i+1}{N_i})f(x+e_i) \\
& \qquad\quad -\frac{x_i}{N_i}f(x-e_i) - (1 - \frac{x_i}{N_i})f(x) \\
& \quad = \quad \frac{x_i}{N_i}[f(x) - f(x-e_i)] + \frac{1}{N_i}f(x) \\
& \qquad\quad +(1 - \frac{x_i+1}{N_i})[f(x+e_i) - f(x)] - \frac{1}{N_i}f(x) \\
& \quad \geq \quad 0.
\end{aligned}
$$

Here the inequality follows from $f \in \mathcal{I}^N(i)$. Hence $\mathcal{T}^N_{ID(i)} : \mathcal{I}^N(i) \to \mathcal{I}^N(i)$, moreover for $j \neq i$ trivially we have $\mathcal{T}^N_{ID(i)} : \mathcal{I}^N(j) \to \mathcal{I}^N(j)$ as well. This implies $\mathcal{T}^N_{ID(i)} f \in \mathcal{I}^N$, and, since $1 \leq i \leq K$ was chosen arbitrarily, we have $\mathcal{T}^N_{ID} : \mathcal{I}^N \to \mathcal{I}^N$.

We continue with the proof of $\mathcal{T}^N_{ID} : \mathcal{C}x^N \to \mathcal{C}x^N$. Suppose, that $f \in \mathcal{C}x^N$. Let $1 \leq i \leq K$ be arbitrary. Choose $x$ with $x, x + 2e_i \in \mathbf{S}^N$. Then

$$\begin{aligned}
&\mathcal{T}^N_{ID(i)} f(x+2e_i) - 2\mathcal{T}^N_{ID(i)} f(x+e_i) + \mathcal{T}^N_{ID(i)} f(x) \\
&= \frac{x_i+2}{N_i} f(x+e_i) + (1 - \frac{x_i+2}{N_i}) f(x+2e_i) \\
&\quad -2\frac{x_i+1}{N_i} f(x) - 2(1 - \frac{x_i+1}{N_i}) f(x+e_i) \\
&\quad +\frac{x_i}{N_i} f(x-e_i) + (1 - \frac{x_i}{N_i}) f(x) \\
&= \frac{x_i}{N_i}[f(x+e_i) - 2f(x) + f(x-e_i)] + \frac{2}{N_i}(f(x+e_i) - f(x)) \\
&\quad +(1 - \frac{x_i+2}{N_i})[f(x+2e_i) - 2f(x+e_i) + f(x)] - \frac{2}{N_i}(f(x+e_i) - f(x)) \\
&\geq 0.
\end{aligned}$$

The inequality comes from $f \in \mathcal{C}x^N(i)$. We may conclude that $\mathcal{T}^N_{ID(i)} : \mathcal{C}x^N(i) \to \mathcal{C}x^N(i)$. Further $\mathcal{T}^N_{ID(i)} : \mathcal{C}x^N(j) \to \mathcal{C}x^N(j)$ is trivial. Hence $\mathcal{T}^N_{ID(i)} f \in \mathcal{C}x^N$. Moreover, since $i$ was arbitrary, also $\mathcal{T}^N_{ID} : \mathcal{C}x^N \to \mathcal{C}x^N$.

Next we will show $\mathcal{T}^N_{ID} : \mathcal{S}uper^N \to \mathcal{S}uper^N$. Suppose $f \in \mathcal{S}uper^N$. Let $1 \leq i \neq j \leq K$. Then we have for $x$ with $x, x + e_i + e_j \in \mathbf{S}^N$ that

$$\begin{aligned}
&\mathcal{T}^N_{ID(i)} f(x+e_i+e_j) - \mathcal{T}^N_{ID(i)} f(x+e_i) - \mathcal{T}^N_{ID(i)} f(x+e_j) + \mathcal{T}^N_{ID(i)} f(x) \\
&= \frac{x_i}{N_i}[f(x+e_j) - f(x) - f(x-e_i+e_j) + f(x-e_i)] \\
&\quad +\frac{1}{N_i}(f(x+e_j) - f(x)) \\
&\quad +(1 - \frac{x_i+1}{N_i})[f(x+e_i+e_j) - f(x+e_i) - f(x+e_j) + f(x)] \\
&\quad -\frac{1}{N_i}(f(x+e_j) - f(x)) \\
&\geq 0.
\end{aligned}$$

The inequality follows from $f \in \mathcal{S}uper^N(i,j)$. Hence $\mathcal{T}^N_{ID(i)} f \in \mathcal{S}uper^N(i,j)$. It easily follows, that $\mathcal{T}^N_{ID(i)} : \mathcal{S}uper^N \to \mathcal{S}uper^N$. Then also $\mathcal{T}^N_{ID} : \mathcal{S}uper^N \to \mathcal{S}uper^N$. $\square$

*Proof of Proposition 5.3.3.* First we prove $\mathcal{T}_C : \mathcal{I}^N \to \mathcal{I}^N$. Let $1 \leq i \leq K$. Then for $x$ with $x, x+e_i \in \mathbf{S}^N$ it holds that

$$\begin{aligned} &\mathcal{T}_C f(x+e_i) - \mathcal{T}_C f(x) \\ &\quad = \sum_{j=1}^K c_j x_j + c_i + f(x+e_i) - \Big(\sum_{j=1}^K c_j x_j + f(x)\Big) \\ &\quad = [f(x+e_i) - f(x)] + c_i \\ &\quad \geq 0. \end{aligned}$$

The inequality follows from the assumption that $f \in \mathcal{I}^N(i)$ and $c_i \geq 0$. Hence $\mathcal{T}_C f \in \mathcal{I}^N(i)$. It follows that $\mathcal{T}_C : \mathcal{I}^N \to \mathcal{I}^N$.

Consider the propagation of convexity. To this end, let $1 \leq i \leq K$. For $x$ with $x, x+2e_i \in \mathbf{S}^N$ it holds that

$$\begin{aligned} &\mathcal{T}_C f(x+2e_i) - 2\mathcal{T}_C f(x+e_i) + \mathcal{T}_C f(x) \\ &\quad = \sum_{j=1}^K c_j x_j + 2c_i + f(x+2e_i) \\ &\qquad -2\Big(\sum_{j=1}^K c_j x_j + c_i + f(x+e_i)\Big) \\ &\qquad +\sum_{j=1}^K c_j x_j + f(x) \\ &\quad = f(x+2e_i) - 2f(x+e_i) + f(x) \\ &\quad \geq 0. \end{aligned}$$

The inequality follows directly from the assumption that $f \in \mathcal{C}x^N(i)$. We conclude that $\mathcal{T}_C f \in \mathcal{C}x^N(i)$. This implies $\mathcal{T}_C : \mathcal{C}x^N \to \mathcal{C}x^N$.

For the propagation of supermodularity $\mathcal{T}_C : \mathcal{S}uper^N \to \mathcal{S}uper^N$, let $1 \leq i \neq j \leq K$. Then for $x$ such that $x, x+e_i+e_j \in \mathbf{S}^N$ it holds that

$$\begin{aligned} &\mathcal{T}_C f(x+e_i+e_j) - \mathcal{T}_C f(x+e_i) - \mathcal{T}_C f(x+e_j) + \mathcal{T}_C f(x) \\ &\quad = \sum_{k=1}^K c_k x_k + c_i + c_j + f(x+e_i+e_j) - \Big(\sum_{k=1}^K c_k x_k + c_i + f(x+e_i)\Big) \\ &\qquad -\Big(\sum_{k=1}^K c_k x_k + c_j + f(x+e_j)\Big) + \sum_{k=1}^K c_k x_k + f(x) \end{aligned}$$

$$
\begin{aligned}
&= f(x+e_i+e_j) - f(x+e_i) - f(x+e_j) + f(x) \\
&\geq 0.
\end{aligned}
$$

The inequality follows from $f \in \mathcal{S}uper^N(i,j)$. Hence, $\mathcal{T}_C f \in \mathcal{S}uper^N(i,j)$ for any $1 \leq i \neq j \leq K$, and so $\mathcal{T}_C : \mathcal{S}uper^N \to \mathcal{S}uper^N$. $\square$

*Proof of Proposition 5.3.4.* First we show $\mathcal{T}^N_{CID} : \mathcal{BD}^N \to \mathcal{BD}^N$. It is necessary to take the combination of more operators, because $\mathcal{T}_C$ alone does not propagate $\mathcal{BD}^N$. First, we derive the following inequalities for the increasing departure operators. Let $f \in \mathcal{BD}^N$, let $1 \leq i \leq K$. Then, for $x$ with $x, x+e_i \in \mathbf{S}^N$ we have

$$
\begin{aligned}
&\mathcal{T}^N_{ID(i)} f(x+e_i) - \mathcal{T}^N_{ID(i)} f(x) \\
&\qquad = \frac{x_i}{N_i}[f(x) - f(x-e_i)] + \frac{1}{N_i} f(x) \\
&\qquad + (1 - \frac{x_i+1}{N_i})[f(x+e_i) - f(x)] - \frac{1}{N_i} f(x) \\
&\qquad \leq (1 - \frac{1}{N_i})\frac{c_i}{\beta_i}.
\end{aligned}
\tag{5.5}
$$

The inequality follows from $f \in \mathcal{BD}^N(i)$. Furthermore, for $j \neq i$ trivially

$$
\mathcal{T}^N_{ID(j)} f(x+e_i) - \mathcal{T}^N_{ID(j)} f(x) \leq \frac{c_i}{\beta_i}. \tag{5.6}
$$

Hence for the operator $\mathcal{T}^N_{CID}$ we obtain

$$
\begin{aligned}
&\mathcal{T}^N_{CID} f(x+e_i) - \mathcal{T}^N_{CID} f(x) \\
&\qquad = \beta_N^{-1} \sum_{j \neq i} \beta_j N_j \left(\mathcal{T}^N_{ID(j)} f(x+e_i) - \mathcal{T}^N_{ID(j)} f(x)\right) \\
&\qquad\qquad + \beta_N^{-1} \beta_i N_i \left(\mathcal{T}^N_{ID(i)} f(x+e_i) - \mathcal{T}^N_{ID(i)} f(x)\right) \\
&\qquad\qquad + \sum_j c_j x_j + c_i - \sum_j c_j x_j \\
&\qquad \leq \beta_N^{-1} \Big( \sum_{j \neq i} \beta_j N_j \frac{c_i}{\beta_i} + \beta_i N_i (1 - \frac{1}{N_i}) \frac{c_i}{\beta_i} + c_i \Big) \\
&\qquad = \beta_N^{-1} \left( \beta_N \frac{c_i}{\beta_i} - c_i + c_i \right) \\
&\qquad = \frac{c_i}{\beta_i}.
\end{aligned}
$$

Here the inequality follows from Inequalities (5.5) and (5.6). This yields $\mathcal{T}_{CID}^N f \in \mathcal{BD}^N(i)$, for all $i$. From this the propagation of $\mathcal{BD}^N$ through $\mathcal{T}_{CID}^N$ follows.

Proof of part $ii$). Suppose for all $1 \le i < K$, it holds that $c_i \ge c_{i+1}$, $c_i\mu_i \ge c_{i+1}\mu_{i+1}$, $c_i\mu_i/\beta_i \ge c_{i+1}\mu_{i+1}/\beta_{i+1}$. We will prove that $\mathcal{T}_{CID}^N : \mathcal{I}^N \cap w\mathcal{UI}^N \cap \mathcal{S}uper^N \cap \mathcal{BD}^N \to w\mathcal{UI}^N$. Let $f \in \mathcal{I}^N \cap w\mathcal{UI}^N \cap \mathcal{S}uper^N \cap \mathcal{BD}^N$, take $1 \le i < K$, let $x$ be such that $x, x+e_i+e_{i+1} \in \mathbf{S}^N$. First observe that for $j \neq i, i+1$, we have

$$\begin{aligned}
&\beta_j N_j \big(\mu_i(\mathcal{T}_{ID(j)}^N f(x+e_i+e_{i+1}) - \mathcal{T}_{ID(j)}^N f(x+e_{i+1})) \\
&\quad - \quad \mu_{i+1}(\mathcal{T}_{ID(j)}^N f(x+e_i+e_{i+1}) - \mathcal{T}_{ID(j)}^N f(x+e_i))\big) \ge 0.
\end{aligned} \tag{5.7}$$

Next, we get the following inequality for $\mathcal{T}_{ID(i)}^N$

$$\begin{aligned}
&\beta_i N_i \big(\mu_i(\mathcal{T}_{ID(i)}^N f(x+e_i+e_{i+1}) - \mathcal{T}_{ID(i)}^N f(x+e_{i+1})) \\
&\qquad - \mu_{i+1}(\mathcal{T}_{ID(i)}^N f(x+e_i+e_{i+1}) - \mathcal{T}_{ID(i)}^N f(x+e_i))\big) \\
&= \beta_i\mu_i\big((x_i+1)f(x+e_{i+1}) + (N_i - x_i - 1)f(x+e_i+e_{i+1}) \\
&\qquad - x_i f(x-e_i+e_{i+1}) - (N_i - x_i)f(x+e_{i+1})\big) \\
&\quad - \beta_i\mu_{i+1}\big((x_i+1)f(x+e_{i+1}) + (N_i - x_i - 1)f(x+e_i+e_{i+1}) \\
&\qquad - (x_i+1)f(x) - (N_i - x_i - 1)f(x+e_i)\big) \\
&= \beta_i x_i[\mu_i(f(x+e_{i+1}) - f(x-e_i+e_{i+1})) - \mu_{i+1}(f(x+e_{i+1}) - f(x))] \\
&\quad + \beta_i(N_i - x_i)\big[\mu_i(f(x+e_i+e_{i+1}) - f(x+e_{i+1})) \\
&\qquad\qquad - \mu_{i+1}(f(x+e_i+e_{i+1}) - f(x+e_i))\big] \\
&\quad + \beta_i\big((\mu_{i+1} - \mu_i)(f(x+e_i+e_{i+1}) - f(x+e_{i+1})) - \mu_{i+1}(f(x+e_i) - f(x))\big) \\
&\ge \beta_i\big((\mu_{i+1} - \mu_i)(f(x+e_i+e_{i+1}) - f(x+e_{i+1})) - \mu_{i+1}(f(x+e_i) - f(x))\big).
\end{aligned} \tag{5.8}$$

The inequality is due to $f \in w\mathcal{UI}^N$. For $\mathcal{T}_{ID(i+1)}^N$ it holds that

$$\begin{aligned}
&\beta_{i+1}N_{i+1}\big(\mu_i(\mathcal{T}_{ID(i+1)}^N f(x+e_i+e_{i+1}) - \mathcal{T}_{ID(i+1)}^N f(x+e_{i+1})) \\
&\qquad -\mu_{i+1}(\mathcal{T}_{ID(i+1)}^N f(x+e_i+e_{i+1}) - \mathcal{T}_{ID(i+1)}^N f(x+e_i))\big) \\
&= \beta_{i+1}\mu_i\big((x_{i+1}+1)f(x+e_i) + (N_{i+1} - x_{i+1} - 1)f(x+e_i+e_{i+1}) \\
&\qquad -(x_{i+1}+1)f(x) - (N_{i+1} - x_{i+1} - 1)f(x+e_{i+1})\big) \\
&\quad -\beta_{i+1}\mu_{i+1}\big((x_{i+1}+1)f(x+e_i) + (N_{i+1} - x_{i+1} - 1)f(x+e_i+e_{i+1})
\end{aligned}$$

$$
\begin{aligned}
& \qquad -x_{i+1}f(x+e_i-e_{i+1})-(N_{i+1}-x_{i+1})f(x+e_i)\big) \\
= \;& \beta_{i+1}x_{i+1}[\mu_i(f(x+e_i)-f(x))-\mu_{i+1}(f(x+e_i)-f(x+e_i-e_{i+1}))] \\
& +\beta_{i+1}(N_{i+1}-x_{i+1}-1)\big[\mu_i(f(x+e_i+e_{i+1})-f(x+e_{i+1})) \\
& \qquad\qquad\qquad -\mu_{i+1}(f(x+e_i+e_{i+1})-f(x+e_i))\big] \\
& \qquad +\beta_{i+1}\mu_i(f(x+e_i)-f(x)) \\
\geq \;& \beta_{i+1}\mu_i(f(x+e_i)-f(x)).
\end{aligned} \tag{5.9}
$$

The inequality follows from $w\mathcal{UI}^N$. Combining the three Inequalities (5.7), (5.8) and (5.9) above gives

$$
\begin{aligned}
& \beta_N\big(\mu_i(\mathcal{T}_{ID}^N f(x+e_i+e_{i+1})-\mathcal{T}_{ID}^N f(x+e_{i+1})) \\
& \qquad -\mu_{i+1}(\mathcal{T}_{ID}^N f(x+e_i+e_{i+1})-\mathcal{T}_{ID}^N f(x+e_i))\big) \\
& \geq \beta_i\big((\mu_{i+1}-\mu_i)(f(x+e_i+e_{i+1})-f(x+e_{i+1}))-\mu_{i+1}(f(x+e_i)-f(x))\big) \\
& \quad +\beta_{i+1}\mu_i(f(x+e_i)-f(x)).
\end{aligned}
$$

Hence, we have that

$$
\begin{aligned}
& \beta_N\big(\mu_i(\mathcal{T}_{CID}^N f(x+e_i+e_{i+1})-\mathcal{T}_{CID}^N f(x+e_{i+1})) \\
& \qquad -\mu_{i+1}(\mathcal{T}_{CID}^N f(x+e_i+e_{i+1})-\mathcal{T}_{CID}^N f(x+e_i))\big) \\
& \geq \beta_i\big((\mu_{i+1}-\mu_i)(f(x+e_i+e_{i+1})-f(x+e_{i+1}))-\mu_{i+1}(f(x+e_i)-f(x))\big) \\
& \quad +\beta_{i+1}\mu_i(f(x+e_i)-f(x))+\mu_i c_i-\mu_{i+1}c_{i+1} =: L_i.
\end{aligned} \tag{5.10}
$$

We wish to argue that that $L_i$ is nonnegative. To this end we will make 4 case distinctions with respect to the parameters.

1. Suppose $\mu_i \leq \mu_{i+1}$ and $\beta_i \leq \beta_{i+1}$, then

$$
\begin{aligned}
L_i &= \beta_i\big((\mu_{i+1}-\mu_i)(f(x+e_i+e_{i+1})-f(x+e_{i+1})) \\
&\qquad\qquad -\mu_{i+1}(f(x+e_i)-f(x))\big) \\
&\quad +\beta_{i+1}\mu_i(f(x+e_i)-f(x))+\mu_i c_i-\mu_{i+1}c_{i+1} \\
&\geq \beta_i(\mu_{i+1}-\mu_i)[f(x+e_i+e_{i+1})-f(x+e_i)-f(x+e_{i+1})+f(x)] \\
&\quad +\mu_i(\beta_{i+1}-\beta_i)[f(x+e_i)-f(x)] \\
&\geq 0.
\end{aligned}
$$

The first inequality is due to $c_i\mu_i \geq c_{i+1}\mu_{i+1}$. The second inequality follows from $\mu_i \leq \mu_{i+1}$ and $f \in \mathcal{S}uper^N(i,j)$, together with $\beta_i \leq \beta_{i+1}$ and $f \in \mathcal{I}^N(i)$.

2. Suppose $\mu_i \leq \mu_{i+1}$ and $\beta_i \geq \beta_{i+1}$, then we have

$$\begin{aligned}
L_i = {} & \beta_i\big((\mu_{i+1}-\mu_i)(f(x+e_i+e_{i+1})-f(x+e_{i+1})) \\
& \qquad\qquad\qquad\qquad - \mu_{i+1}(f(x+e_i)-f(x))\big) \\
& + \beta_{i+1}\mu_i(f(x+e_i)-f(x)) + \mu_i c_i - \mu_{i+1}c_{i+1} \\
= {} & \beta_i(\mu_{i+1}-\mu_i)[f(x+e_i+e_{i+1})-f(x+e_i)-f(x+e_{i+1})+f(x)] \\
& - \mu_i(\beta_i-\beta_{i+1})[f(x+e_i)-f(x)] + \mu_i c_i - \mu_{i+1}c_{i+1} \\
\geq {} & - \mu_i(\beta_i-\beta_{i+1})[f(x+e_i)-f(x)] + \mu_i c_i - \mu_{i+1}c_{i+1} \\
\geq {} & - \mu_i(\beta_i-\beta_{i+1})\frac{c_i}{\beta_i} + \mu_i c_i - \mu_{i+1}c_{i+1} \\
= {} & - \mu_i c_i + \beta_{i+1}\frac{\mu_i c_i}{\beta_i} + \mu_i c_i - \mu_{i+1}c_{i+1} \\
\geq {} & - \mu_i c_i + \beta_{i+1}\frac{\mu_{i+1} c_{i+1}}{\beta_{i+1}} + \mu_i c_i - \mu_{i+1}c_{i+1} \\
= {} & 0.
\end{aligned}$$

The first inequality is due to $f \in \mathcal{S}uper^N(i,j)$ together with $\mu_i \leq \mu_{i+1}$. The second inequality comes from $\beta_i \geq \beta_{i+1}$ and $f \in \mathcal{BD}^N(i)$. The last inequality is due to $c_i\mu_i/\beta_i \geq c_{i+1}\mu_{i+1}/\beta_{i+1}$.

3. Next we assume that $\mu_i \geq \mu_{i+1}$, $\mu_i/\beta_i \geq \mu_{i+1}/\beta_{i+1}$, then

$$\begin{aligned}
L_i = {} & \beta_i\big((\mu_{i+1}-\mu_i)(f(x+e_i+e_{i+1})-f(x+e_{i+1})) \\
& \qquad\qquad\qquad\qquad - \mu_{i+1}(f(x+e_i)-f(x))\big) \\
& + \beta_{i+1}\mu_i(f(x+e_i)-f(x)) + \mu_i c_i - \mu_{i+1}c_{i+1} \\
= {} & (\beta_{i+1}\mu_i - \beta_i\mu_{i+1})[f(x+e_i)-f(x)] \\
& - \beta_i(\mu_i-\mu_{i+1})[f(x+e_i+e_{i+1})-f(x+e_{i+1})] + \mu_i c_i - \mu_{i+1}c_{i+1} \\
\geq {} & - c_i(\mu_i-\mu_{i+1}) + \mu_i c_i - \mu_{i+1}c_{i+1} \\
= {} & \mu_{i+1}(c_i - c_{i+1}) \\
\geq {} & 0.
\end{aligned}$$

The first inequality follows from $\mu_i/\beta_i \geq \mu_{i+1}/\beta_{i+1}$ and $f \in \mathcal{I}^N(i)$, together with $\mu_i \geq \mu_{i+1}$ and $f \in \mathcal{BD}^N(i)$. The second inequality comes from $c_i \geq c_{i+1}$.

4. Finally assume $\mu_i \geq \mu_{i+1}$ and $\mu_i/\beta_i \leq \mu_{i+1}/\beta_{i+1}$, then

$$\begin{aligned}
L_i = \beta_i\big(&(\mu_{i+1} - \mu_i)(f(x+e_i+e_{i+1}) - f(x+e_{i+1})) \\
&\qquad - \mu_{i+1}(f(x+e_i) - f(x))\big) \\
&+ \beta_{i+1}\mu_i(f(x+e_i) - f(x)) + \mu_i c_i - \mu_{i+1}c_{i+1} \\
= {}& (\beta_i\mu_{i+1} - \beta_{i+1}\mu_i)[f(x+e_i+e_{i+1}) - f(x+e_i) \\
&\qquad - f(x+e_{i+1}) + f(x)] + \mu_i c_i - \mu_{i+1}c_{i+1} \\
&- (\beta_i - \beta_{i+1})\mu_i[f(x+e_i+e_{i+1}) - f(x+e_{i+1})] \\
\geq {}& - (\beta_i - \beta_{i+1})\mu_i\frac{c_i}{\beta_i} + \mu_i c_i - \mu_{i+1}c_{i+1} \\
= {}& + \beta_{i+1}\frac{c_i\mu_i}{\beta_i} - c_{i+1}\mu_{i+1} \\
\geq {}& + \beta_{i+1}\frac{c_{i+1}\mu_{i+1}}{\beta_{i+1}} - c_{i+1}\mu_{i+1} \\
= {}& 0.
\end{aligned}$$

The first inequality follows from $\mu_i/\beta_i \leq \mu_{i+1}/\beta_{i+1}$ and $f \in \mathcal{S}uper^N(i,j)$, together with $\beta_i \geq \beta_{i+1}$ and $f \in \mathcal{BD}^N(i)$. The third inequality is due to $c_i\mu_i/\beta_i \geq c_{i+1}\mu_{i+1}/\beta_{i+1}$.

So for any $1 \leq i < K$, we have $\mathcal{T}_{CID}^N f \in w\mathcal{UI}^N(i)$. Hence, we conclude that $\mathcal{T}_{CID}^N : \mathcal{I}^N \cap w\mathcal{UI}^N \cap \mathcal{S}uper^N \cap \mathcal{BD}^N \to w\mathcal{UI}^N$. □

*Proof of Proposition 5.3.5.* Before starting with the proofs, we have the following remark. By Lemma 5.2.3 we have that $v^\alpha \in w\mathcal{UI} \cap \mathcal{I}$ implies the smallest index policy to be optimal. The same holds true if $f \in w\mathcal{UI}^N \cap \mathcal{I}^N$. Then for $x \in \mathbf{S}^N$, the action that chooses the smallest index minimises the $\mathcal{T}_{MS}$ operator. We will use this several times below.

First we prove that $\mathcal{T}_{MS} : \mathcal{I}^N \cap w\mathcal{UI}^N \to \mathcal{I}^N$. Assume that $f \in \mathcal{I}^N \cap w\mathcal{UI}^N$. Let $1 \leq i \leq K$ be arbitrary. It suffices to show that $\mathcal{T}_{MS}f \in \mathcal{I}^N(i)$. First suppose $x = 0$. Then we have

$$\begin{aligned}
&\mathcal{T}_{MS}f(x+e_i) - \mathcal{T}_{MS}f(x) \\
&\qquad = \quad \mathcal{T}_{MS}f(e_i) - \mathcal{T}_{MS}f(0) \\
&\qquad = \quad \frac{1}{\mu}\min_j\{\mu_j f(0) + (\mu - \mu_j)f(e_i)\} - f(0) \\
&\qquad = \quad \frac{\mu - \mu_i}{\mu}[f(e_i) - f(0)] \\
&\qquad \geq \quad 0.
\end{aligned}$$

The optimal policy is non-idling because $f \in \mathcal{I}^N$, hence in state $e_i$ the system is minimised by serving station $i$, the only non-empty queue. In the second term the system is empty, so nobody is served. The inequality follows from $f \in \mathcal{I}^N(i)$.

Next suppose that $x \neq 0$, with $x, x+e_i \in \mathbf{S}^N$. Let $j^*$ be the optimal action in state $x$. If $j^* \leq i$ then $w\mathcal{UI}^N$ implies that this is also the optimal action in state $x+e_i$. The inequality follows straightforwardly. If $j^* > i$ then $w\mathcal{UI}^N$ implies that it is optimal to serve station $i$ in state $x+e_i$. We obtain

$$\begin{aligned}
&\mathcal{T}_{MS}f(x+e_i) - \mathcal{T}_{MS}f(x) \\
&\quad= \frac{1}{\mu}\min_j\{\mu_j f((x+e_i-e_j)^+) + (\mu-\mu_j)f(x+e_i)\} \\
&\qquad -\frac{1}{\mu}\min_j\{\mu_j f((x-e_j)^+) + (\mu-\mu_j)f(x)\} \\
&\quad= \frac{1}{\mu}\Big(\mu_i f(x) + (\mu-\mu_i)f(x+e_i) \\
&\qquad -\mu_{j^*}f(x-e_{j^*}) - (\mu-\mu_{j^*})f(x)\Big) \\
&\quad= \frac{1}{\mu}\Big(\mu_{j^*}[f(x)-f(x-e_{j^*})] - (\mu_{j^*}-\mu_i)f(x) \\
&\qquad +(\mu-\mu_i)[f(x+e_i)-f(x)] + (\mu_{j^*}-\mu_i)f(x)\Big) \\
&\quad\geq 0.
\end{aligned}$$

The last inequality follows from $f \in \mathcal{I}^N$. We conclude that $\mathcal{T}_{MS}f \in \mathcal{I}^N(i)$.

We continue by proving $\mathcal{T}_{MS} : \mathcal{I}^N \cap w\mathcal{UI}^N \to w\mathcal{UI}^N$. Suppose that $f \in \mathcal{I}^N \cap w\mathcal{UI}^N$, so that the smallest index policy is optimal. Let $1 \leq i < K$ be arbitrary. Let $x$ be such that $x, x+e_i+e_{i+1} \in \mathbf{S}^N$. Let $j^*$ be the optimal action in state $x+e_{i+1}$. Suppose that $j^* \leq i$, then the smallest index policy implies $j^*$ to be optimal in the states $x+e_i$, $x+e_{i+1}$ and $x+e_i+e_{i+1}$ as well. The propagation of $w\mathcal{UI}^N$ is trivial. Suppose that $j^* > i$. Then action $i$ is optimal in states $x+e_i$ and $x+e_i+e_{i+1}$, while action $i+1$ is optimal in state $x+e_{i+1}$. We get

$$\begin{aligned}
&\mu\Big(\mu_i(\mathcal{T}_{MS}f(x+e_i+e_{i+1}) - \mathcal{T}_{MS}f(x+e_{i+1})) \\
&\qquad -\mu_{i+1}(\mathcal{T}_{MS}f(x+e_i+e_{i+1}) - \mathcal{T}_{MS}f(x+e_i))\Big) \\
&\quad= \mu_i\Big(\min_j\{\mu_j f((x+e_i+e_{i+1}-e_j)^+) + (\mu-\mu_j)f(x+e_i+e_{i+1})\}
\end{aligned}$$

$$
\begin{aligned}
& \quad - \min_j \{\mu_j f((x+e_{i+1}-e_j)^+) + (\mu-\mu_j) f(x+e_{i+1})\}\Big) \\
& -\mu_{i+1}\Big(\min_j \{\mu_j f((x+e_i+e_{i+1}-e_j)^+) + (\mu-\mu_j) f(x+e_i+e_{i+1})\} \\
& \quad - \min_j \{\mu_j f((x+e_i-e_j)^+) + (\mu-\mu_j) f(x+e_i)\}\Big) \\
= \; & \mu_i(\mu_i f(x+e_{i+1}) + (\mu-\mu_i) f(x+e_i+e_{i+1})) \\
& -\mu_i(\mu_{i+1} f(x) + (\mu-\mu_{i+1}) f(x+e_{i+1})) \\
& -\mu_{i+1}(\mu_i f(x+e_{i+1}) + (\mu-\mu_i) f(x+e_i+e_{i+1})) \\
& +\mu_{i+1}(\mu_i f(x) + (\mu-\mu_i) f(x+e_i)) \\
= \; & (\mu-\mu_i)[\mu_i(f(x+e_i+e_{i+1}) - f(x+e_{i+1})) \\
& \qquad -\mu_{i+1}(f(x+e_i+e_{i+1}) - f(x+e_i))] \\
\geq \; & 0.
\end{aligned}
$$

The inequality follows from the assumption $f \in w\mathcal{UI}^N(i)$. We conclude that $\mathcal{T}_{MS} : \mathcal{I}^N \cap w\mathcal{UI}^N \to w\mathcal{UI}^N(i)$, implying $\mathcal{T}_{MS} : \mathcal{I}^N \cap w\mathcal{UI}^N \to w\mathcal{UI}^N$.

Proof of $\mathcal{T}_{MS} : \mathcal{I}^N \cap w\mathcal{UI}^N \cap \mathcal{C}x^N \cap \mathcal{S}uper^N \to \mathcal{C}x^N \cap \mathcal{S}uper^N$. To this end, assume $f \in \mathcal{I}^N \cap w\mathcal{UI}^N \cap \mathcal{C}x^N \cap \mathcal{S}uper^N$. Let $1 \leq i \neq j \leq K$, in particular assume $i < j$. First suppose that $x = 0$. Then it is optimal to serve station $i$ in states $x + e_i$ and $x + e_i + e_j$ and it is optimal to take action $j$ in state $x + e_j$. Therefore, it holds that

$$
\begin{aligned}
& \mu\big(\mathcal{T}_{MS} f(x+e_i+e_j) - \mathcal{T}_{MS} f(x+e_i) - \mathcal{T}_{MS} f(x+e_j) + \mathcal{T}_{MS} f(x)\big) \\
& \quad = \; \min_k \{\mu_k f((x+e_i+e_j-e_k)^+) + (\mu-\mu_k) f(x+e_i+e_j)\} \\
& \qquad - \min_k \{\mu_k f((x+e_i-e_k)^+) + (\mu-\mu_k) f(x+e_i)\} \\
& \qquad - \min_k \{\mu_k f((x+e_j-e_k)^+) + (\mu-\mu_k) f(x+e_j)\} \\
& \qquad + \min_k \{\mu_k f((x-e_k)^+) + (\mu-\mu_k) f(x)\} \\
& \quad = \; \mu_i f(e_j) + (\mu-\mu_i) f(e_i+e_j) \\
& \qquad -\mu_i f(0) - (\mu-\mu_i) f(e_i) \\
& \qquad -\mu_j f(0) - (\mu-\mu_j) f(e_j) \\
& \qquad +\mu f(0) \\
& \quad = \; (\mu-\mu_i)[f(e_i+e_j) - f(e_i) - f(e_j) + f(0)] \\
& \qquad +\mu_j[f(e_j) - f(0)] \\
& \quad \geq \; 0.
\end{aligned}
$$

The inequality follows from $f \in \mathcal{I}^N(j) \cap \mathcal{S}uper^N(i,j)$.

Next, suppose $x \neq 0$, with $x, x+e_i+e_j \in \mathbf{S}^N$. Let the optimal action in state $x$ be $j^*$. We will make three case distinctions. First suppose $j^* \leq i$, then $f \in \mathcal{I}^N \cap w\mathcal{UI}^N$ implies that $j^*$ is also optimal in $x+e_i$, $x+e_j$ and $x+e_i+e_j$. The same action is optimal in every state, which implies that $\mathcal{S}uper^N(i,j)$ is propagated trivially. If $i < j^* \leq j$, then $f \in \mathcal{I}^N \cap w\mathcal{UI}^N$ implies $j^*$ to be also optimal in $x+e_j$, and action $i$ is optimal in $x+e_i$ and $x+e_i+e_j$. We obtain

$$\begin{aligned}
&\mu\big(\mathcal{T}_{MS}f(x+e_i+e_j) - \mathcal{T}_{MS}f(x+e_i) - \mathcal{T}_{MS}f(x+e_j) + \mathcal{T}_{MS}f(x)\big)\\
&\quad = \quad \mu_i f(x+e_j) + (\mu-\mu_i)f(x+e_i+e_j)\\
&\quad - \quad \mu_i f(x) - (\mu-\mu_i)f(x+e_i)\\
&\quad - \quad \mu_{j^*} f(x+e_j-e_{j^*}) - (\mu-\mu_{j^*})f(x+e_j)\\
&\quad + \quad \mu_{j^*} f(x-e_{j^*}) + (\mu-\mu_{j^*})f(x)\\
&\quad = \quad \mu_{j^*}[f(x+e_j) - f(x) - f(x+e_j-e_{j^*}) + f(x-e_{j^*})]\\
&\qquad\qquad\qquad -(\mu_{j^*}-\mu_i)(f(x+e_j)-f(x))\\
&\quad + \quad \mu_i[f(x+e_i+e_j) - f(x+e_i) - f(x+e_j) + f(x)]\\
&\qquad\qquad\qquad +(\mu_{j^*}-\mu_i)(f(x+e_j)-f(x))\\
&\quad \geq \quad 0
\end{aligned}$$

The inequality follows from $f \in \mathcal{S}uper^N \cap \mathcal{C}x^N$. If $j^* > j$, then $f \in \mathcal{I}^N \cap w\mathcal{UI}^N$ implies that action $i$ is optimal in $x+e_i$ and $x+e_i+e_j$. Serving station $j$ is optimal in state $x+e_j$. If we choose the suboptimal action $j^*$ instead, this makes the expression only smaller. Then we are in the same situation as above, for which we already derived that the expression is nonnegative. Hence, $\mathcal{T}_{MS}f \in \mathcal{S}uper^N(i,j)$. We conclude that $\mathcal{T}_{MS}f \in \mathcal{S}uper^N$ as well.

This only leaves to prove that $\mathcal{T}_{MS}f \in \mathcal{C}x^N$. First, consider the case that $x=0$. Then action $i$ is optimal in states $x+e_i$ and $x+2e_i$. Hence, we have

$$\begin{aligned}
&\mu\big(\mathcal{T}_{MS}f(x+2e_i) - 2\mathcal{T}_{MS}f(x+e_i) + \mathcal{T}_{MS}f(x)\big)\\
&\quad = \quad \min_j\{\mu_j f((x+2e_i-e_j)^+) + (\mu-\mu_j)f(x+2e_i)\}\\
&\qquad\quad -2\min_j\{\mu_j f((x+e_i-e_j)^+) + (\mu-\mu_j)f(x+e_i)\}\\
&\qquad\quad +\min_j\{\mu_j f((x-e_j)^+) + (\mu-\mu_j)f(x)\}\\
&\quad = \quad \mu_i f(e_i) + (\mu-\mu_i)f(2e_i)\\
&\qquad\quad -2\mu_i f(0) - 2(\mu-\mu_i)f(e_i)\\
&\qquad\quad +\mu f(0)
\end{aligned}$$

$$
\begin{aligned}
&= (\mu - \mu_i)[f(2e_i) - 2f(e_i) + f(0)] \\
&\quad +\mu_i[f(e_i) - f(0)] \\
&\geq 0.
\end{aligned}
$$

The inequality follows from $f \in \mathcal{I}^N(i) \cap \mathcal{C}x^N(i)$.

Next, consider the case that $x \neq 0$, with $x, x+2e_i \in \mathbf{S}^N$. Let $j^*$ be the the optimal action in state $x$. If $j^* \leq i$, then in states $x+e_i$ and $x+2e_i$ the optimal actions are equal to $j^*$ as well. Propagation is trivial in that case. If $j^* > i$, then the optimal action in states $x+e_i$ and $x+2e_i$ is action $i$. We obtain the following inequality

$$
\begin{aligned}
&\mu\big(\mathcal{T}_{MS}f(x+2e_i) - 2\mathcal{T}_{MS}f(x+e_i) + \mathcal{T}_{MS}f(x)\big) \\
&\quad = \min_j\{\mu_j f((x+2e_i-e_j)^+) + (\mu-\mu_j)f(x+2e_i)\} \\
&\qquad -2\min_j\{\mu_j f((x+e_i-e_j)^+) + (\mu-\mu_j)f(x+e_i)\} \\
&\qquad + \min_j\{\mu_j f((x-e_j)^+) + (\mu-\mu_j)f(x)\} \\
&\quad = \mu_i f(x+e_i) + (\mu-\mu_i)f(x+2e_i) \\
&\qquad -2\mu_i f(x) - 2(\mu-\mu_i)f(x+e_i) \\
&\qquad +\mu_{j^*} f(x-e_{j^*}) + (\mu-\mu_{j^*})f(x) \\
&\quad \geq \mu_i f(x+e_i) + (\mu-\mu_i)f(x+2e_i) \\
&\qquad -\mu_i f(x) - (\mu-\mu_i)f(x+e_i) \\
&\qquad -\mu_{j^*} f(x+e_i-e_{j^*}) - (\mu-\mu_{j^*})f(x+e_i) \\
&\qquad +\mu_{j^*} f(x-e_{j^*}) + (\mu-\mu_{j^*})f(x) \\
&\quad = \mu_{j^*}[f(x+e_i) - f(x) - f(x+e_i-e_{j^*}) + f(x-e_{j^*})] \\
&\qquad +(\mu-\mu_i)[f(x+2e_i) - 2f(x+e_i) + f(x)] \\
&\quad \geq 0.
\end{aligned}
$$

The first inequality is due to $f \in w\mathcal{UI}^N$, the second comes from $f \in \mathcal{S}uper^N \cap \mathcal{C}x^N$. Hence, $\mathcal{T}_{MS}f \in \mathcal{C}x^N(i)$. We conclude that $\mathcal{T}_{MS} : \mathcal{I}^N \cap w\mathcal{UI}^N \cap \mathcal{C}x^N \cap \mathcal{S}uper^N \to \mathcal{C}x^N \cap \mathcal{S}uper^N$.

We prove *ii*), and assume $c_i\mu_i/\beta_i \geq c_{i+1}\mu_{i+1}/\beta_{i+1}$, for all $1 \leq i < K$. We will prove that $\mathcal{T}_{MS} : \mathcal{I}^N \cap w\mathcal{UI}^N \cap \mathcal{BD}^N \to \mathcal{BD}^N$. Let $f \in \mathcal{I}^N \cap w\mathcal{UI}^N \cap \mathcal{BD}^N$, and let $1 \leq i \leq K$ be arbitrary. Again we make two case distinctions. First suppose $x = 0$, then we get

$$\begin{aligned}
&\mathcal{T}_{MS}f(x+e_i) - \mathcal{T}_{MS}f(x) \\
&\quad = (1-\alpha)\frac{\mu-\mu_i}{\mu}[f(e_i)-f(0)] \\
&\quad \leq [f(e_i)-f(0)] \\
&\quad \leq \frac{c_i}{\beta_i}.
\end{aligned}$$

For the second inequality we use that $f \in \mathcal{BD}^N(i)$. Next suppose $x \neq 0$. Let $j^*$ be the optimal action in state $x$. If $j^* \leq i$ then as a result of $f \in \mathcal{I}^N \cap w\mathcal{UI}^N$ the optimal actions in states $x$ and $x+e_i$ are equal, namely $j^*$. As a consequence the propagation is trivial. If $j^* > i$ then the optimal actions are not equal, because $w\mathcal{UI}^N$ implies that in state $x+e_i$ it is optimal to serve state $i$. We obtain

$$\begin{aligned}
&\mathcal{T}_{MS}f(x+e_i) - \mathcal{T}_{MS}f(x) \\
&\quad = \frac{1}{\mu}\Big(\mu_{j^*}[f(x)-f(x-e_{j^*})] - (\mu_{j^*}-\mu_i)f(x) \\
&\quad\quad +(\mu-\mu_i)[f(x+e_i)-f(x)] + (\mu_{j^*}-\mu_i)f(x)\Big) \\
&\quad \leq \frac{1}{\mu}\Big(\mu_{j^*}\frac{c_{j^*}}{\beta_{j^*}} + (\mu-\mu_i)\frac{c_i}{\beta_i}\Big) \\
&\quad \leq \frac{1}{\mu}\Big(\mu_i\frac{c_i}{\beta_i} + (\mu-\mu_i)\frac{c_i}{\beta_i}\Big) \\
&\quad = \frac{c_i}{\beta_i}.
\end{aligned}$$

The first inequality follows from $f \in \mathcal{BD}^N$, the second follows from$c_i\mu_i/\beta_i \geq c_{j^*}\mu_{j^*}/\beta_{j^*}$ for $i < j^*$. We conclude that $\mathcal{T}_{MS}f \in \mathcal{BD}^N(i)$ and thus

$$\mathcal{T}_{MS} : \mathcal{I}^N \cap w\mathcal{UI}^N \cap \mathcal{BD}^N \to \mathcal{BD}^N.$$

□

*Proof of Proposition 5.3.6.* It follows directly that $\mathcal{T}_{UNIF} : (\mathcal{I}^N)^3 \to \mathcal{I}^N$, $(w\mathcal{UI}^N)^3 \to w\mathcal{UI}^N$, $(\mathcal{C}x^N)^3 \to \mathcal{C}x^N$, $(\mathcal{S}uper^N)^3 \to \mathcal{S}uper^N$, since convex combinations of nonnegative terms are nonnegative.

The propagation $\mathcal{T}_{UNIF} : (\mathcal{BD}^N)^3 \to \mathcal{BD}^N$ is also straightforward. We have that $\lambda + \beta_N + \mu = 1$, hence if $f_1, f_2, f_3 \in \mathcal{BD}^N$, then

$$\mathcal{T}_{UNIF}(f_1, f_2, f_3) := \lambda f_1 + \beta_N f_2 + \mu f_3 \in \mathcal{BD}^N.$$

$\square$

*Proof of Proposition 5.3.7.* Recall that $\mathcal{T}^{\bar{\alpha}}_{DISC} f = (1-\bar{\alpha})f$. Further, clearly $\mathcal{T}^{\bar{\alpha}}_{DISC} : w\mathcal{UI}^N \rightarrow,\ \mathcal{I}^N \rightarrow \mathcal{I}^N,\ \mathcal{C}x^N \rightarrow \mathcal{C}x^N, \mathcal{S}uper^N \rightarrow \mathcal{S}uper^N$.

Suppose, that $f \in \mathcal{BD}^N$. Let $1 \leq i \leq K$ be arbitrary, let $x$ be such that $x, x+e_i \in \mathbf{S}^N$. Then

$$\begin{aligned}
&\mathcal{T}^{\bar{\alpha}}_{DISC} f(x+e_i) - \mathcal{T}^{\bar{\alpha}}_{DISC} f(x) \\
&\quad = \quad (1-\bar{\alpha})[f(x+e_i) - f(x)] \\
&\quad \leq \quad (1-\bar{\alpha})\frac{c_i}{\beta_i} \\
&\quad \leq \quad \frac{c_i}{\beta_i}.
\end{aligned}$$

The first inequality is due to $f \in \mathcal{BD}^N(i)$. This implies $\mathcal{T}^{\bar{\alpha}}_{DISC} f \in \mathcal{BD}^N(i)$. Hence, $\mathcal{T}^{\bar{\alpha}}_{DISC} : \mathcal{BD}^N \rightarrow \mathcal{BD}^N$.

# 6 Truncations of a tandem queue

This chapter is based on work in progress by Blok and Spieksma.

## 6.1 Introduction

In this chapter, we discuss a model illustrating how several truncations can have a different effect on the structural results, and thus on the optimal policy. Chapter 2 introduces a tandem queue that can be modelled as a discrete time Markov decision process with a countable state space (see Figure 2.1 in Section 2.2.2). In this case, a truncation is not needed to make the process uniformisable. However, if one would try to compute the optimal policy numerically, a finite state space truncation is necessary. If a straightforward truncation is executed, the outcome is not even close to the true optimal policy, as is evidenced by Figure 2.2.

In Section 2.2.2 several claims have been made without detailed proofs. The current chapter aims to provide the missing details. In Section 6.2 we introduce the optimal policy and the relation to structural properties. In Section 6.3 we prove the structural properties via value iteration. First we do this for the untruncated model, then for the model with Smoothed Rate Truncation. In Section 6.4 we give some final remarks on different truncations.

## 6.2 Optimal policy and properties of value function

Central to this chapter is the following theorem.

**Theorem 6.2.1.** *Let $\alpha \in (0, 1)$. Suppose that $\mu_1(1 - p_1) \geq \mu_2$, then AP1 is $\alpha$-discount optimal, that is, give full priority to station 1.*

We will prove this theorem lateron. Moreover, we show that this theorem also holds true for a collection of MDPs that are perturbed by a smoothed rate truncation as described below. The lemma below reduces the proof of Theorem 6.2.1 to deriving a structural property of the value function. To this end, define

$$\begin{aligned}\mathcal{FP}(p_1) := \big\{ f : \mathbf{S} \to \mathbb{R} \,|\, \mu_2 f(x-e_2) - \mu_1(1-p_1) f(x-e_1) \\ -\mu_1 p_1 f(x-e_1+e_2) + (\mu_1 - \mu_2) f(x) \geq 0, x \in \mathbf{S}, x_1, x_2 > 0 \big\}.\end{aligned}$$

**Lemma 6.2.2.** *Let $v^\alpha$ be the discounted value function for the tandem queue. Then $v^\alpha \in \mathcal{FP}(p_1)$ implies that AP1 is $\alpha$-discount optimal.*

*Proof.* The discounted cost optimality equation is given by

$$\begin{aligned}u(x) = x_1 + x_2 + (1-\alpha)\Big(\lambda u(x+e_i) \\ + \min\big\{\mu_1 p_1 u(x-e_1+e_2) + \mu_1(1-p_1) u(x-e_1) + \mu_2 u(x), \\ \mu_1 u(x) + \mu_2 u(x-e_2)\big\}\Big). \qquad (6.1)\end{aligned}$$

Due to Section 2.2.1, $v^\alpha$ is the unique solution to Eq. (6.1) and any minimising policy is optimal. Now, $v^\alpha \in \mathcal{FP}(p_1)$ implies that for all $x$ with $x_1, x_2 > 0$, the first argument of the minimisation is less than or equal to the second argument. This implies that AP1 is a minimising policy in Eq. (6.1), thus AP1 is optimal. □

## 6.3 Dynamic programming

With value iteration it can be shown that $\mu_1(1-p_1) \geq \mu_2$ implies $v^\alpha \in \mathcal{FP}(p_1)$. We will do this by means of the event based dynamic programming.

### 6.3.1 No Truncation

Define the operators $\mathcal{T}_C, \mathcal{T}^\alpha_{DISC}, \mathcal{T}_{UNIF}, \mathcal{T}_{A(1)}, \mathcal{T}_{MS(p_1)}$ for $x \in \mathbf{S}$ by

$$\begin{aligned}\mathcal{T}_C f(x) &:= f(x) + x_1 + x_2;\\ \mathcal{T}^\alpha_{DISC} f &:= (1-\alpha) f;\\ \mathcal{T}_{UNIF}(f_1, f_2) &:= \lambda f_1 + (\mu_1+\mu_2) f_2;\\ \mathcal{T}_{A(1)} f(x) &:= f(x+e_1);\end{aligned}$$

$$\mathcal{T}_{MS(p_1)} f(x) :=$$
$$\frac{1}{\mu_1+\mu_2}\begin{cases}\min\big\{\mu_1 p_1 f(x-e_1+e_2) + \mu_1(1-p_1) f(x-e_1)) + \mu_2 f(x), \\ \qquad \mu_1 f(x) + \mu_2 f((x-e_2)^+)\big\} & x_1 > 0;\\ \mu_1 f(x) + \mu_2 f((x-e_2)^+) & x_1 = 0.\end{cases}$$

Then the value iteration scheme is given by $v_0 \equiv 0$ and for $x \in \mathbf{S}$

$$\begin{aligned} v_{n+1}(x) &= \mathcal{T}_C\big(\mathcal{T}^{\alpha}_{DISC}\big(\mathcal{T}_{UNIF}(\mathcal{T}_{A(1)}v_n(x), \mathcal{T}_{MS(p_1)}v_n(x))\big)\big) \\ &= x_1 + x_2 + (1-\alpha)\Big(\lambda v_n(x+e_i) \\ &\quad + \min\big\{\mu_1 p_1 v_n(x-e_1+e_2) + \mu_1(1-p_1)v_n(x-e_1) + \mu_2 v_n(x), \\ &\qquad\qquad \mu_1 v_n(x) + \mu_2 v_n(x-e_2)\big\}\Big) \end{aligned} \tag{6.2}$$

**Lemma 6.3.1.** *The following propagations hold.*

*a) For all* $\alpha \in (0,1)$

$$\mathcal{T}^{\alpha}_{DISC} : \mathcal{FP}(p_1) \to \mathcal{FP}(p_1),$$

*b)*

$$\mathcal{T}_{UNIF} : \mathcal{FP}(p_1) \times \mathcal{FP}(p_1) \to \mathcal{FP}(p_1),$$

*c)*

$$\mathcal{T}_{A(1)} : \mathcal{FP}(p_1) \to \mathcal{FP}(p_1),$$

*d)*

$$\mathcal{T}_{MS(p_1)} : \mathcal{FP}(p_1) \to \mathcal{FP}(p_1),$$

*e) if moreover* $\mu_1(1-p_1) \geq \mu_2$*, then*

$$\mathcal{T}_C : \mathcal{FP}(p_1) \to \mathcal{FP}(p_1).$$

*Proof.* Suppose $f, f_1, f_2 \in \mathcal{FP}(p_1)$. Let $x \in \mathbf{S}$ with $x_1, x_2 > 0$.

*Proof of a.* Let $\alpha \in (0,1)$, then

$$\begin{aligned} &\mu_2\mathcal{T}^{\alpha}_{DISC}f(x-e_2) - \mu_1(1-p_1)\mathcal{T}^{\alpha}_{DISC}f(x-e_1) \\ &\qquad - \mu_1 p_1\mathcal{T}^{\alpha}_{DISC}f(x-e_1+e_2) + (\mu_1-\mu_2)\mathcal{T}^{\alpha}_{DISC}f(x) \\ &=(1-\alpha)\big[\mu_2 f_1(x-e_2) - \mu_1(1-p_1)f_1(x-e_1) \\ &\qquad\qquad - \mu_1 p_1 f_1(x-e_1+e_2) + (\mu_1-\mu_2)f_1(x)\big] \\ &\geq 0. \end{aligned}$$

The inequality follows from $f \in \mathcal{FP}(p_1)$.

*Proof of b.* We have that

$$\begin{aligned}
&\mu_2\mathcal{T}_{UNIF}(f_1,f_2)(x-e_2)-\mu_1(1-p_1)\mathcal{T}_{UNIF}(f_1,f_2)(x-e_1)\\
&\qquad-\mu_1p_1\mathcal{T}_{UNIF}(f_1,f_2)(x-e_1+e_2)+(\mu_1-\mu_2)\mathcal{T}_{UNIF}(f_1,f_2)(x)\\
&\quad=\lambda\big[\mu_2f_1(x-e_2)-\mu_1(1-p_1)f_1(x-e_1)\\
&\qquad\qquad-\mu_1p_1f_1(x-e_1+e_2)+(\mu_1-\mu_2)f_1(x)\big]\\
&\qquad+(\mu_1+\mu_2)\Big[\mu_2f_2(x-e_2)-\mu_1(1-p_1)f_2(x-e_1)\\
&\qquad\qquad-\mu_1p_1f_2(x-e_1+e_2)+(\mu_1-\mu_2)f_2(x)\Big]\\
&\quad\geq 0.
\end{aligned}$$

The inequality is due to $f_1, f_2 \in \mathcal{FP}(p_1)$.

*Proof of c.* We have that

$$\begin{aligned}
&\mu_2\mathcal{T}_{A(1)}f(x-e_2)-\mu_1(1-p_1)\mathcal{T}_{A(1)}f(x-e_1)\\
&\qquad\qquad-\mu_1p_1\mathcal{T}_{A(1)}f(x-e_1+e_2)+(\mu_1-\mu_2)\mathcal{T}_{A(1)}f(x)\\
&=\mu_2f(x+e_1-e_2)-\mu_1(1-p_1)f(x)-\mu_1p_1f(x+e_2)+(\mu_1-\mu_2)f(x+e_1)\\
&\geq 0.
\end{aligned}$$

*Proof of d.* Let $x_1 > 1$. Then $f \in \mathcal{FP}(p_1)$ implies that action 1 is optimal in states $x-e_2,\ x-e_1,\ x-e_1+e_2$ and $x$. Hence

$$\begin{aligned}
&\mu_2\mathcal{T}_{MS(p_1)}f(x-e_2)-\mu_1(1-p_1)\mathcal{T}_{MS(p_1)}f(x-e_1)\\
&\qquad\qquad-\mu_1p_1\mathcal{T}_{MS(p_1)}f(x-e_1+e_2)+(\mu_1-\mu_2)\mathcal{T}_{MS(p_1)}f(x)\\
&\quad=\mu_1p_1\big[\mu_2f(x-e_1)-\mu_1(1-p_1)f(x-2e_1+e_2)\\
&\qquad\qquad-\mu_1p_1f(x-2e_1+2e_2)+(\mu_1-\mu_2)f(x-e_1+e_2)\big]\\
&\qquad+\mu_1(1-p_1)\big[\mu_2f(x-e_1-e_2)-\mu_1(1-p_1)f(x-2e_1)\\
&\qquad\qquad-\mu_1p_1f(x-2e_1+e_2)+(\mu_1-\mu_2)f(x-e_1)\big]\\
&\qquad+\mu_2\big[\mu_2f(x-2e_2)-\mu_1(1-p_1)f(x-e_1-e_2)\\
&\qquad\qquad-\mu_1p_1f(x-e_1)+(\mu_1-\mu_2)f(x-e_2)\big]\\
&\quad\geq 0.
\end{aligned}$$

The inequality comes from $f \in \mathcal{FP}(p_1)$. Now let $x_1 = 1$. Then action 1 is optimal in $x-e_2$ and $x$, whereas in states $x-e_1$ and $x-e_1+e_2$ only the second station is nonempty. Hence

$$\begin{aligned}
&\mu_2\mathcal{T}_{A(1)}f(x-e_2)-\mu_1(1-p_1)\mathcal{T}_{A(1)}f(x-e_1)-\\
&\qquad\qquad\mu_1p_1\mathcal{T}_{A(1)}f(x-e_1+e_2)+(\mu_1-\mu_2)\mathcal{T}_{A(1)}f(x)\\
&=\mu_2\big(\mu_1p_1f(x-e_1)+\mu_1(1-p_1)f(x-e_1-e_2)+\mu_2f(x-e_2)\big)\\
&\quad-\mu_1(1-p_1)\big(\mu_1f(x-e_1)+\mu_2f(x-e_1-e_2)\big)\\
&\quad-\mu_1p_1\big(\mu_1f(x-e_1+e_2)+\mu_2f(x-e_1)\big)\\
&\quad+(\mu_1-\mu_2)\big(\mu_1p_1f(x-e_1+e_2)+\mu_1(1-p_1)f(x-e_1)+\mu_2f(x)\big)\\
&=\mu_2\big[\mu_2f(x-e_2)-\mu_1(1-p_1)f(x-e_1)-\mu_1p_1f(x-e_1+e_2)+(\mu_1-\mu_2)f(x)\big]\\
&\geq 0.
\end{aligned}$$

The inequality is due to $f\in\mathcal{FP}(p_1)$.

*Proof of e).* Suppose $\mu_1(1-p_1)-\mu_2\geq 0$. Then

$$\begin{aligned}
&\mu_2\mathcal{T}_Cf(x-e_2)-\mu_1(1-p_1)\mathcal{T}_Cf(x-e_1)\\
&\qquad\qquad-\mu_1p_1\mathcal{T}_Cf(x-e_1+e_2)+(\mu_1-\mu_2)\mathcal{T}_Cf(x)\\
&=\mu_2f(x-e_2)-\mu_1(1-p_1)f(x-e_1)-\mu_1p_1f(x-e_1+e_2)+(\mu_1-\mu_2)f(x)\\
&\quad+\mu_1(1-p_1)-\mu_2\\
&\geq 0.
\end{aligned}$$

The inequality follows from $f\in\mathcal{FP}(p_1)$ and $\mu_1(1-p_1)\geq\mu_2$.

□

**Corollary 6.3.2.** *Let $\alpha\in(0,1)$. For the tandem with no truncation, if $\mu_1(1-p_1)\geq\mu_2$, then*

*a) for all $n\geq 0$, $v_n\in\mathcal{FP}(p_1)$;*

*b) $v^\alpha\in\mathcal{FP}(p_1)$.*

*Proof.* Combining Eq. (6.2) and Lemma 6.3.1, the first assertion follows by induction. The second assertion is by convergence of value iteration, which follows from Theorem 2.2.3. The assumptions of the theorem, can be easily checked by putting the drift function $V(x)=e^{\epsilon(x_1+x_2)}$. □

*Proof of Theorem 6.2.1.* Corollary 6.3.2*b)* and Lemma 6.2.2 yield the result. □

### 6.3.2 Smoothed rate truncation

In this section we apply a smoothed rate truncation on the process to make the state space bounded. In the same style as for the nontruncated Markov decision process, we will show that AP1 is an $\alpha$-discount optimal policy. Smoothed rate truncation prescribes to linearly truncate the rates of events that move the system into a higher state. However, we wish to keep the complexity of the analysis as low as possible. It turns out that it is sufficient to truncate the arrival rate as a function of both $x_1 + x_2$, thereby leaving the positive recurrent states to be a triangular set.

The SRT arrival rates in state $x$ will be

$$\lambda_{12}^N(x) := \frac{(N_{12} - x_1 - x_2) \vee 0}{N_{12}} \lambda.$$

This gives a new arrival operator

$$\mathcal{T}_{SA(1)}^{N_{12}} f(x) := \frac{(N_{12} - x_1 - x_2) \vee 0}{N_{12}} f(x + e_1) + \frac{(x_1 + x_2) \wedge N_{12}}{N_{12}} f(x).$$

The new value iteration scheme is then given by $v_0 \equiv 0$ and

$$v_{n+1}(x) = \mathcal{T}_C\big(\mathcal{T}_{DISC}^\alpha\big(\mathcal{T}_{UNIF}(\mathcal{T}_{SA(1)}^{N_{12}} v_n(x), \mathcal{T}_{MS(p_1)} v_n(x))\big)\big). \tag{6.3}$$

It is straightforward to see that

$$\Delta(N_{12}) = \{x \in \mathbf{S} | x_1 + x_2 \le N_{12}\}$$

is a positive recurrent closed class. Therefore we define the property $\mathcal{FP}(p_1)$ now only to hold on $\Delta(N_{12})$

$$\begin{aligned}\mathcal{FP}^{N_{12}}(p_1) := \big\{f : \mathbf{S} \to \mathbb{R} | \mu_2 f(x - e_2) - \mu_1(1 - p_1) f(x - e_1) \\ - \mu_1 p_1 f(x - e_1 + e_2) + (\mu_1 - \mu_2) f(x) \ge 0,\ x \in \Delta(N_{12}), x_1, x_2 > 0\big\}.\end{aligned}$$

For the propagation of $\mathcal{FP}^{N_{12}}(p_1)$ we need an additional property, namely upstream increasing, denoted by $\mathcal{UI}^{N_{12}}(1)$ (cf. Section 7.3). Notice that $\mathcal{UI}^{N_{12}}(1)$ is equal to $\mathcal{FP}^{N_{12}}(p_1)$ for the special case that $\mu_1 = \mu_2$ and $p_1 = 0$. The following lemma gives the propagation results.

**Lemma 6.3.3.** *The following propagations hold.*

*a) For all* $\alpha \in (0, 1)$

$$\mathcal{T}_{DISC}^\alpha : \mathcal{FP}^{N_{12}}(p_1) \to \mathcal{FP}^{N_{12}}(p_1),\ \mathcal{UI}^{N_{12}}(1) \to \mathcal{UI}^{N_{12}}(1),$$

*b)*

$$\mathcal{T}_{UNIF} : \mathcal{FP}^{N_{12}}(p_1) \times \mathcal{FP}^{N_{12}}(p_1) \to \mathcal{FP}^{N_{12}}(p_1),$$
$$\mathcal{UI}^{N_{12}}(1) \times \mathcal{UI}^{N_{12}}(1) \to \mathcal{UI}^{N_{12}}(1),$$

*c)*

$$\mathcal{T}^{N_{12}}_{SA(1)} : \mathcal{FP}^{N_{12}}(p_1) \cap \mathcal{UI}^{N_{12}}(1) \to \mathcal{FP}^{N_{12}}(p_1),\ \mathcal{UI}^{N_{12}}(1) \to \mathcal{UI}^{N_{12}}(1),$$

*d)*

$$\mathcal{T}_{MS(p_1)} : \mathcal{FP}^{N_{12}}(p_1) \to \mathcal{FP}^{N_{12}}(p_1),\ \mathcal{UI}^{N_{12}}(1) \to \mathcal{UI}^{N_{12}}(1),$$

*e)*

$$\mathcal{T}_C : \mathcal{UI}^{N_{12}}(1) \to \mathcal{UI}^{N_{12}}(1),$$

*if moreover* $\mu_1(1-p_1) \geq \mu_2$, *then*

$$\mathcal{T}_C : \mathcal{FP}^{N_{12}}(p_1) \to \mathcal{FP}^{N_{12}}(p_1).$$

*Proof.* First notice that because $\mathcal{UI}^{N_{12}}(1)$ is a special case of $\mathcal{FP}^{N_{12}}(p_1)$ the propagation results for $\mathcal{UI}^{N_{12}}(1)$ follow immediately from the ones for $\mathcal{FP}^{N_{12}}(p_1)$. We will omit them here.

Suppose $f \in \mathcal{FP}^{N_{12}}(p_1)$, let $x \in \Delta(N_{12})$, with $x_1, x_2 > 0$. Then the proofs for Assertions a), b), d) and e) are identical to Lemma 6.3.1). Thus it remains to prove Assertion c. To this end suppose $f \in \mathcal{FP}^{N_{12}}(p_1) \cap \mathcal{UI}^{N_{12}}(1)$, let $x \in \Delta(N_{12})$, with $x_1, x_2 > 0$. Then

$$\begin{aligned}
N_{12}\big(&\mu_2 \mathcal{T}^{N_{12}}_{SA(1)} f(x-e_2) - \mu_1(1-p_1)\mathcal{T}^{N_{12}}_{SA(1)} f(x-e_1) \\
&\qquad - \mu_1 p_1 \mathcal{T}^{N_{12}}_{SA(1)} f(x-e_1+e_2) + (\mu_1-\mu_2)\mathcal{T}^{N_{12}}_{SA(1)} f(x)\big) \\
=&\mu_2\Big((N_{12}-x_1-x_2+1)f(x+e_1-e_2) + (x_1+x_2-1)f(x-e_2)\Big) \\
&- \mu_1(1-p_1)\Big((N_{12}-x_1-x_2+1)f(x) + (x_1+x_2-1)f(x-e_1)\Big) \\
&- \mu_1 p_1\Big((N_{12}-x_1-x_2)f(x+e_2) + (x_1+x_2)f(x-e_1+e_2)\Big) \\
&+ (\mu_1-\mu_2)\Big((N_{12}-x_1-x_2)f(x+e_1) + (x_1+x_2)f(x)\Big) \\
=&(N_{12}-x_1-x_2)\big[\mu_2 f(x+e_1-e_2) - \mu_1(1-p_1)f(x) - \mu_1 p_1 f(x+e_2) \\
&\qquad + (\mu_1-\mu_2)f(x+e_1)\big] \quad + \mu_2 f(x+e_1-e_2) - \mu_1(1-p_1)f(x) \\
&+ (x_1+x_2-1)\big[\mu_2 f(x-e_2) - \mu_1(1-p_1)f(x-e_1) - \mu_1 p_1 f(x-e_1+e_2)
\end{aligned}$$

$$
\begin{aligned}
&\quad + (\mu_1 - \mu_2) f(x)] \quad - \mu_1 p_1 f(x - e_1 + e_2) + (\mu_1 - \mu_2) f(x) \\
&\geq \mu_2 [f(x + e_1 - e_2) - f(x)] + \mu_1 p_1 [f(x) - f(x - e_1 + e_2)] \\
&\geq 0.
\end{aligned}
$$

The first inequality is due to $f \in \mathcal{FP}^{N_{12}}(p_1)$, the second inequality comes from $f \in \mathcal{UI}^{N_{12}}(1)$. □

**Corollary 6.3.4.** *Consider the smoothed rate truncated tandem queue problem. Let $\alpha \in (0,1)$, let $N_{12} \in \mathbb{Z}_+$. If $\mu_1(1 - p_1) \geq \mu_2$, then*

*a) for all $n \geq 0$, $v_n \in \mathcal{FP}^{N_{12}}(p_1)$;*

*b) $v^\alpha \in \mathcal{FP}^{N_{12}}(p_1)$.*

*Proof.* Combining Eq. (6.3) and Lemma 6.3.3 the first assertion follows by induction. The second assertion is by convergence of value iteration, which follows from Theorem 2.2.3. □

**Corollary 6.3.5.** *Consider the smoothed rate truncated tandem queue problem. Let $\alpha \in (0,1)$, let $N_{12} \in \mathbb{Z}_+$. If $\mu_1(1 - p_1) \geq \mu_2$, then AP1 is $\alpha$-discount optimal.*

*Proof.* Due to Corollary 6.3.4 *b)*, $v^\alpha \in \mathcal{FP}^{N_{12}}(p_1)$. Then by a reasoning similar to the proof of Lemma 6.2.2, it follows that AP1 is $\alpha$-discount optimal. □

## 6.4 Conclusion

A few remarks are at place at the end of this chapter. In the first place, a warning that monotonicity properties can be destroyed under naive truncations. Secondly, it should be noted that smoothed rate truncation can be used to circumvent this, and this shows that SRT can be useful for uniformisable MDPs as well. Further, Figure 2.3 suggests that rectangular SRT is applicable as well, emphasising that SRT is not a uniquely defined method. Different choices are possible, perhaps one prevaling over the other. In this case, the triangular SRT seemed the easiest. Finally it should be noted that a state space truncation with a triangular state space would have worked as well, provided the right choice would have been made on the boundary: namely, simply blocking arriving customers, if the sum of the two queues equals the truncation value.

# 7 Event Based dynamic programming

This chapter is based on work in progress by Blok and Spieksma.

## 7.1 Introduction

In Chapters 4 and 5 different models are analysed and structural properties are derived. The analysis is performed on truncated (and uniformisable) versions of the original unbounded-rate MDP. The Chapters 2 and 3 provide verification methods to deduce that results for the approximating MDPs converge to the one with unbounded rates.

This analysis is done by means of event based dynamic programming. The $(n+1)$-th step value iteration function can be composed from operators representing events in the Markov decision process applied to the $n$-th step value iteration function. If it can be shown that the event operators propagate certain properties, then it follows that the $(n+1)$-th step value iteration function itself possesses these properties provides the $n$-step value iteration function has. If the appropriate conditions are fulfilled, then value iteration converges and we can conclude that the value function has these properties as well.

The monograph of Koole [45] is a comprehensive treatment of event based dynamic programming. Naturally – since value iteration is a discrete time algorithm – the monograph restricts to uniformisable Markov decision processes.

However, by truncation events arise that have adapted rates compared to the standard events. This gives rise to new event operators. In Chapters 4, 5 and 6 some new operators have been introduced in combination with propagation results regarding the properties that are relevant for the respective models, cf. also [16].

In this chapter we aim to give a systematic overview of the operators that are associated with truncated Markov decision processes. Furthermore, we will provide the obtained propagation results. We would like to point out, the obtained propagation results in the previous chapters that seem to be very model specific, have not been not listed here.

Generally speaking, the truncations can be of two different types.

1. Truncations that prevent the system to move beyond a finite bound. This makes the number of states that are positive recurrent finite. Because the essential state space (the positive recurrent class of states) is finite, the resulting transition rates are bounded. Hence, this makes the system uniformisable. The best way to do this, seems to be smoothed rate truncation (see Chapters 5 and 6). A straightforward state space truncation might work as well, but often breaks the properties of interest. Because the essential state space is finite, it is sufficient to derive properties on a finite state space.

2. Truncations that directly bound the rates (see Chapter 4). Such a perturbation does not have an effect on whether states are communicating or not. This implies that – for any given policy – the set of essential states is invariant under the perturbation. So for these perturbations it is required to show the properties on the entire state space.

Within one model, a combination of these two is possible as well. Usually when a smoothed rate truncation is performed on one or more events, then the rates of other events are not allowed to increase outside the essential state space.

The rest of this chapter contains the following. In Section 7.2 we introduce the operators related to unbounded rates. In Section 7.3 we define the structural properties of interest. The propagation results are presented in Section 7.4.

## 7.2 Operators related to truncations

Let $K \in \mathbb{N}$, define

$$\mathbf{S} := \{x \in \mathbb{Z}^K | x_i \geq 0, \text{ for } 1 \leq i \leq K\},$$

$$\mathcal{N} := \{N \in \mathbb{Z}^K | N_i \geq 1, \text{ for } 1 \leq i \leq K\}.$$

For $N \in \mathcal{N}$ define

$$\mathbf{S}^N := \{x \in \mathbb{Z}^K | 0 \leq x_i \leq N_i, \text{ for } 1 \leq i \leq K\}.$$

In the rest of this section, let $i \in \{1, \ldots, K\}$, $N \in \mathcal{N}$, define the following operators.

**SRT Arrivals operator** Define $\mathcal{T}^N_{SA(i)}$ by

$$\mathcal{T}^N_{SA(i)} f(x) := \frac{(N_i - x_i) \vee 0}{N_i} f(x + e_i) + \frac{x_i \wedge N_i}{N_i} f(x).$$

This operator is equal to the finite source operator $\mathcal{T}^B_{FS(i)}$ as in [45]. However, the propagation results are more extensive. Moreover the operator has a different interpretation, which opens new applications for it. Results for this operator stem from [16] and Chapter 5.

For $\lambda_i \geq 0$, $i = 1, \ldots, K$, $\sum_{i=1}^K \lambda_i = \lambda > 0$, we define the Total SRT Arrivals operator as

$$\mathcal{T}^N_{SA} = \frac{1}{\lambda} \sum_{i=1}^K \lambda_i \mathcal{T}^N_{SA(i)}.$$

**SRT Controlled Arrivals operator** For $c_i, c'_i \in \mathbb{R}$, let $\mathcal{T}^N_{SCA(i)}$ be given by

$$\begin{aligned}\mathcal{T}^N_{SCA(i)} f(x) := \frac{(N_i - x_i) \vee 0}{N_i} &\min\{c_i + f(x), c'_i + f(x + e_i)\} \\ &+ \frac{x_i \wedge N_i}{N_i} (c_i + f(x)).\end{aligned}$$

This operator models control for customer arrivals. It has been introduced in [16].

For $\lambda_i \geq 0$, $i = \{1, \ldots, K\}$, $\sum_{i=1}^K \lambda_i = \lambda > 0$, we define the Total SRT Controlled Arrivals operator as

$$\mathcal{T}^N_{SCA} = \lambda^{-1} \sum_{i=1}^K \lambda_i \mathcal{T}^N_{SCA(i)}.$$

**SRT Double Arrivals operator** Let $j \in 1, \ldots, K$, $j \neq i$, $N_i = N_j = N_{ij}$, define $\mathcal{T}^N_{SDA(i,j)}$ by

$$\mathcal{T}^N_{SDA(i,j)} f(x) := \begin{cases} \left(1 - \frac{x_i + x_j}{N_{ij}}\right) f(x + e_i + e_j) + \frac{x_j}{N_{ij}} f(x + e_i) + \frac{x_i}{N_{ij}} f(x + e_j) \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{if } x_i + x_j \leq N_{ij}, \\ \left(1 - \left(1 - \frac{x_i}{N_{ij}}\right)^+ - \left(1 - \frac{x_j}{N_{ij}}\right)^+\right) f(x) \\ \qquad + \left(1 - \frac{x_i}{N_{ij}}\right)^+ f(x + e_i) + \left(1 - \frac{x_j}{N_{ij}}\right)^+ f(x + e_j) \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{if } x_i + x_j > N_{ij}. \end{cases}$$

This operator is from [16], where a rather complicated truncation, with two layers of different truncations, was necessary. It models the event of batch arrivals of size 2. It might be possible to extend this truncation to batch arrivals of larger size. However, the number of layers grows as the size of the batch, which makes the analysis too complicated already for batch size 3.

**SRT Transfers operator** Let $i < K$, define $\mathcal{T}^N_{ST(i)}$ by

$$\mathcal{T}^N_{ST(i)} f(x) := \begin{cases} \frac{(N_{i+1}-x_{i+1})\vee 0}{N_{i+1}} f(x - e_i + e_{i+1}) + \frac{x_{i+1}\wedge N_{i+1}}{N_{i+1}} f(x) & x_i > 0, \\ f(x) & x_i = 0. \end{cases}$$

This operator is from [16]. It models transfers from queue $i$ to the next queue $i+1$.

For $\mu_i > 0$, $i = 1, \dots, K-1$ and $\sum_{i=1}^{K-1} \mu_i = \mu$, we define the Total SRT Transfers operator by

$$\mathcal{T}^N_{ST} = \mu^{-1} \sum_{i=1}^{K-1} \mu_i \mathcal{T}^N_{ST(i)}.$$

**SRT Increasing Transfers operator** For $i < K$, define $\mathcal{T}^N_{SIT(i)}$ by

$$\begin{aligned} \mathcal{T}^N_{SIT(i)} f(x) := {} & \frac{x_i \wedge N_i}{N_i} \frac{(N_{i+1} - x_{i+1}) \vee 0}{N_{i+1}} f(x - e_i + e_{i+1}) \\ & + \frac{x_{i+1} \wedge N_{i+1}}{N_{i+1}} \frac{(N_i - x_i) \vee 0}{N_i} f(x + e_i - e_{i+1}) \\ & + \Big(1 - \frac{x_i \wedge N_i}{N_i} \frac{(N_{i+1} - x_{i+1}) \vee 0}{N_{i+1}} - \frac{x_{i+1} \wedge N_{i+1}}{N_{i+1}} \frac{(N_i - x_i) \vee 0}{N_i}\Big) f(x). \end{aligned}$$

The origin of this operator is the analysis in [11], where increasing transfers were present in opposite directions. In the event based setting it has been introduced in [16]. This operator may represent increasing transfers from station $i$ to $i+1$ only. Take $N_i = \epsilon N_j$ with $\epsilon$ small. It follows that for $x \in \mathbf{S}^N$,

$$\mathcal{T}^N_{SIT(i)} f(x) \underset{\epsilon \to 0}{\to} \frac{x_i}{N_i} \frac{N_{i+1} - x_{i+1}}{N_{i+1}} f(x - e_i + e_{i+1}) + \frac{N_i - x_i}{N_i} \frac{x_{i+1}}{N_{i+1}} f(x).$$

**Increasing Departures operator** Define $\mathcal{T}^N_{ID(i)}$ as follows

$$\mathcal{T}^N_{ID(i)} f(x) := \frac{x_i \wedge N_i}{N_i} f(x - e_i) + \frac{(N_i - x_i) \vee 0}{N_i} f(x).$$

Propagations for this operator have been obtained in [16] and Chapter 5. This operator can be seen as a special case of $\mathcal{T}_{D(i)}$ from [45]. $\mathcal{T}^N_{ID(i)}$ may represent an $N_i$-server system, an approximation of an infinite server system or an approximation of customer abandonments. The latter two represent models with unbounded rates. This operator does not bound the state space, but only the transition rates.

For $\beta_i > 0$, $i = 1, \ldots, K$, $\sum_{i=1}^K \beta_i N_i = \beta_N$ define the Total Increasing Departures operator by

$$\mathcal{T}^N_{ID} = \beta_N^{-1} \sum_{i=1}^K \beta_i N_i \mathcal{T}^N_{ID(i)}.$$

**Increasing Idle/Off operator** Let $i < K$, $c \geq 0$, define $\mathcal{T}^N_{I/O(i+1)}$ by

$$\mathcal{T}^N_{I/O(i+1)} f(x) := \frac{x_{i+1} \wedge N_{i+1}}{N_{i+1}} \min\{f(x + e_i - e_{i+1}), f(x - e_{i+1}) + c\} + \frac{(N_{i+1} - x_{i+1}) \vee 0}{N_{i+1}} f(x).$$

This operator is from [1] and has been treated in Chapter 4 as well. It models the choice upon a service completion between turning a server off or leaving it idle. The rates are the same as in $\mathcal{T}^N_{ID(i+1)}$, so the operator bounds the transition rates, not the state space.

## 7.3 Properties

The monotonicity properties that are relevant in the models that we studied are non-decreasingness, convexity, supermodularity, superconvexity, upstream increasingness and downstream increasingness. The first three concepts are standard. For the latter three we follow the notation of [45]. We define these properties through collections of functions that have the respective properties. As mentioned before, it is useful to consider both properties that hold on the entire state space, as well as properties that hold on the restriction to $\mathbf{S}^N$. Moreover, we consider one property both in one dimension and in more dimensions.

Let $i, j \in \{1, \ldots, K\}$ with $i \neq j$, let $N \in \mathcal{N}$. Define the following collections.

### Non-decreasing (Increasing) functions

$$\begin{aligned}
\mathcal{I}(i) &:= \{f : \mathbf{S} \to \mathbb{R} | f(x+e_i) - f(x) \geq 0, \text{ for } x, x+e_i \in \mathbf{S}\};\\
\mathcal{I} &:= \bigcap_{1 \leq i \leq K} \mathcal{I}(i);\\
\mathcal{I}^N(i) &:= \{f : \mathbf{S} \to \mathbb{R} | f(x+e_i) - f(x) \geq 0, \text{ for } x, x+e_i \in \mathbf{S}^N\};\\
\mathcal{I}^N &:= \bigcap_{1 \leq i \leq K} \mathcal{I}^N(i).
\end{aligned}$$

### Convex functions

$$\begin{aligned}
\mathcal{C}x(i) &:= \{f : \mathbf{S} \to \mathbb{R} | f(x+2e_i) - 2f(x+e_i) + f(x) \geq 0,\\
&\qquad \text{for } x, x+2e_i \in \mathbf{S}\};\\
\mathcal{C}x &:= \bigcap_{1 \leq i \leq K} \mathcal{C}x(i);\\
\mathcal{C}x^N(i) &:= \{f : \mathbf{S} \to \mathbb{R} | f(x+2e_i) - 2f(x+e_i) + f(x) \geq 0,\\
&\qquad \text{for } x, x+2e_i \in \mathbf{S}^N\};\\
\mathcal{C}x^N &:= \bigcap_{1 \leq i \leq K} \mathcal{C}x^N(i).
\end{aligned}$$

### Supermodular functions

$$\begin{aligned}
\mathcal{S}uper(i,j) &:= \{f : \mathbf{S} \to \mathbb{R} | f(x+e_i+e_j) - f(x+e_i) - f(x+e_j) + f(x) \geq 0,\\
&\qquad \text{for } x, x+e_i+e_j \in \mathbf{S}\};\\
\mathcal{S}uper &:= \bigcap_{1 \leq i \neq j \leq K} \mathcal{S}uper(i,j);\\
\mathcal{S}uper^N(i,j) &:= \{f : \mathbf{S} \to \mathbb{R} | f(x+e_i+e_j) - f(x+e_i) - f(x+e_j) + f(x) \geq 0,\\
&\qquad \text{for } x, x+e_i+e_j \in \mathbf{S}^N\};\\
\mathcal{S}uper^N &:= \bigcap_{1 \leq i \neq j \leq K} \mathcal{S}uper^N(i,j).
\end{aligned}$$

### Superconvex functions

$$\begin{aligned}
\mathcal{S}uper\mathcal{C}(i,j) &:= \{f : \mathbf{S} \to \mathbb{R} | f(x+2e_i) - f(x+e_i)\\
&\qquad - f(x+e_i+e_j) + f(x+e_j) \geq 0, \text{for } x, x+e_i+e_j \in \mathbf{S}\}.
\end{aligned}$$

**Downstream Increasing functions**

$$\begin{aligned}\mathcal{DI}^N(i) &:= \{f : \mathbf{S} \to \mathbb{R} | f(x + e_{i+1}) - f(x + e_i) \geq 0, \\ &\qquad\qquad\qquad\qquad \text{for } x, x + e_i, x + e_{i+1} \in \mathbf{S}^N\}; \\ \mathcal{DI}^N &:= \bigcap_{1 \leq i < K} \mathcal{DI}^N(i).\end{aligned}$$

**Upstream Increasing functions**

$$\begin{aligned}\mathcal{UI}^N(i) &:= \{f : \mathbf{S} \to \mathbb{R} | f(x + e_{i+1}) - f(x + e_i) \leq 0, \\ &\qquad\qquad\qquad\qquad \text{for } x, x + e_i, x + e_{i+1} \in \mathbf{S}^N\}; \\ \mathcal{UI}^N &:= \bigcap_{1 \leq i < K} \mathcal{UI}^N(i).\end{aligned}$$

## 7.4 Propagations

Because smoothed rate truncation is designed to reduce the essential states to $\mathbf{S}^N$, it is sufficient to derive propagation results for the SRT operators restricted to $\mathbf{S}^N$.

### 7.4.1 SRT Arrivals

**Theorem 7.4.1.** *Let $N \in \mathcal{N}$, then for $i \in \{1, \ldots, K\}$*

*1.*
$$\mathcal{T}^N_{SA(i)} : \mathcal{I}^N(i) \to \mathcal{I}^N(i),$$

*2.*
$$\mathcal{T}^N_{SA(i)} : \mathcal{C}x^N(i) \to \mathcal{C}x^N(i),$$

*3.*
$$\mathcal{T}^N_{SA(i)} : \mathcal{S}uper^N(i,j) \to \mathcal{S}uper^N(i,j),$$

*4. if $i < K$, then*

$$\mathcal{T}^N_{SA(i)} : \mathcal{DI}^N(i) \cap \mathcal{I}^N(i+1) \to \mathcal{DI}^N(i).$$

*Proof.* Propagations 1, 2 and 3 have been proven in Proposition 5.3.1. Propagation 4 is from [16, p.37]. For completeness we give the proof of 4 here. Let $i \in \{1, \ldots, K-1\}$, suppose that $f \in \mathcal{DI}^N(i) \cap \mathcal{I}^N(i+1)$. Let $x$ be such that $x + e_i, x + e_{i+1} \in \mathbf{S}^N$, then $x_i \leq N_i - 1$. It holds that

$$\begin{aligned}
&N_i(\mathcal{T}^N_{SA(i)} f(x + e_{i+1}) - \mathcal{T}^N_{SA(i)} f(x + e_i)) \\
&\quad = \quad (N_i - x_i) f(x + e_i + e_{i+1}) + x_i f(x + e_{i+1}) \\
&\quad\quad\quad -(N_i - x_i - 1) f(x + 2e_i) - (x_i + 1) f(x + e_i) \\
&\quad = \quad (N_i - x_i - 1)[f(x + e_i + e_{i+1}) - f(x + 2e_i)] \\
&\quad\quad\quad + x_i[f(x + e_{i+1}) - f(x + e_i)] \\
&\quad\quad\quad + f(x + e_i + e_{i+1}) - f(x + e_i) \\
&\quad \geq \quad 0.
\end{aligned}$$

For the inequality we used that $f \in \mathcal{DI}^N(i) \cap \mathcal{I}^N(i+1)$. □

**Corollary 7.4.2.** *Consider Let $N \in \mathcal{N}$, then*

1. $$\mathcal{T}^N_{SA} : \mathcal{I}^N \to \mathcal{I}^N,$$

2. $$\mathcal{T}^N_{SA} : \mathcal{C}x^N \to \mathcal{C}x^N,$$

3. $$\mathcal{T}^N_{SA} : \mathcal{S}uper^N \to \mathcal{S}uper^N,$$

4. $$\mathcal{T}^N_{SA} : \mathcal{DI}^N \cap \mathcal{I}^N \to \mathcal{DI}^N.$$

*Proof.* This is a direct consequence from Theorem 7.4.1. □

### 7.4.2 SRT Controlled Arrivals

**Theorem 7.4.3.** *Let $N \in \mathcal{N}$, then for $i \in \{1, \ldots, K\}$,*

1. $$\mathcal{T}^N_{SCA(i)} : \mathcal{I}^N(i) \to \mathcal{I}^N(i),$$

2. $$\mathcal{T}^N_{SCA(i)} : \mathcal{C}x^N(i) \to \mathcal{C}x^N(i),$$

*3.*
$$\mathcal{T}^N_{SCA(i)} : \mathcal{S}uper^N(i,j) \to \mathcal{S}uper^N(i,j),$$

*4. if $i \geq 2$, then*
$$\mathcal{T}^N_{SA(i)} : \mathcal{DI}^N(i-1) \to \mathcal{DI}^N(i-1).$$

*Proof.* The results are from [16, p.51], for completeness we give them here. Let $i \in \{1, \ldots, K\}$.

*Proof of 1.* Suppose $f \in \mathcal{I}^N(i)$. Let $x \in \mathbf{S}^N$, so that $x_i \leq N_i - 1$. It holds that

$$\begin{aligned}
&N_i(\mathcal{T}^N_{SCA(i)} f(x+e_i) - \mathcal{T}^N_{SCA(i)} f(x)) \\
=&(N_i - x_i - 1)\min\{c_i + f(x+e_i), c_i' + f(x+2e_i)\} + (x_i+1)(c_i + f(x+e_i)) \\
&- (N_i - x_i)\min\{c_i + f(x), c_i' + f(x+e_i)\} - x_i(c_i + f(x)) \\
=&(N_i - x_i - 1)[\underbrace{\min\{c_i + f(x+e_i), c_i' + f(x+2e_i)\}}_{(1)} \\
&\qquad - \underbrace{\min\{c_i + f(x), c_i' + f(x+e_i)\}}_{(2)}] - \underbrace{\min\{c_i + f(x), c_i' + f(x+e_i)\}}_{(3)} \\
&+ x_i[f(x+e_i) - f(x)] + (c_i + f(x+e_i)) \\
\geq&0.
\end{aligned}$$

To prove that this is greater than or equal to zero, we have to make some case distinctions. If the minimum of (1) is $c_i + f(x+e_i)$ (reject), then in (2) we also choose reject (this can only make the expression smaller). Then the terms inside the square brackets are nonnegative by $f \in \mathcal{I}^N(i)$. If the minimum of (1) is accept, then in (2) choose accept and the inequality between the square brackets also follows from $f \in \mathcal{I}^N(i)$. In (3) always choose reject, the nonnegativity of the remaining terms again follows from $f \in \mathcal{I}^N(i)$.

*Proof of 2.* Suppose $f \in \mathcal{C}x^N(i)$, let $x$ be such that $x + 2e_i \in \mathbf{S}^N$. Then $x_i \leq N_i - 2$, and we obtain

$$\begin{aligned}
&N_i(T^S_{CA(i)} f(x+2e_i) - 2T^S_{CA(i)} f(x+e_i) + T^S_{CA(i)} f(x)) \\
=&(N_i - x_i - 2)\min\{c_i + f(x+2e_i), c_i' + f(x+3e_i)\} \\
&\qquad + (x_i+2)(c_i + f(x+2e_i)) \\
&- 2(N_i - x_i - 1)\min\{c_i + f(x+e_i), c_i' + f(x+2e_i)\} \\
&\qquad - 2(x_i+1)(c_i + f(x+e_i))
\end{aligned}$$

$$
\begin{aligned}
&+ (N_i - x_i)\min\{c_i + f(x), c_i' + f(x+e_i)\} + x_i(c_i + f(x))\\
=&(N_i - x_i - 2)\\
&\times[\underbrace{\min\{c_i + f(x+2e_i), c_i' + f(x+3e_i)\}}_{(1)} - \underbrace{\min\{c_i + f(x+e_i), c_i' + f(x+2e_i)\}}_{(2)}\\
&- \underbrace{\min\{c_i + f(x+e_i), c_i' + f(x+2e_i)\}}_{(3)} + \underbrace{\min\{c_i + f(x), c_i' + f(x+e_i)\}}_{(4)}]\\
&- 2\underbrace{\min\{c_i + f(x+e_i), c_i' + f(x+2e_i)\}}_{(5)} + 2\underbrace{\min\{c_i + f(x), c_i' + f(x+e_i)\}}_{(6)}\\
&+ x_i[f(x+2e_i) - 2f(x+e_i) + f(x)] + 2f(x+2e_i) - 2f(x+e_i)\\
\geq&0.
\end{aligned}
$$

Again choosing any action in (2), (3) or (5) can only make the expression smaller. Let the actions of (2), (3) and (5) copy the minimisers of (1), (4) and (6) respectively. Case checking shows that the expression can be reduced to a convex expression. The assumption that $f \in \mathcal{C}x^N(i)$ yields the result.

*Proof of 3.* Suppose $f \in \mathcal{S}uper^N(i,j)$, let $x$ such that $x + e_i + e_j \in \mathbf{S}^N$. Hence $x_i \leq N_i - 1$ and thus

$$
\begin{aligned}
&N_i(T^S_{CA(i)}f(x) + T^S_{CA(i)}f(x+e_i+e_j) - T^S_{CA(i)}f(x+e_i) - T^S_{CA(i)}f(x+e_j))\\
=&(N_i - x_i)\min\{c_i + f(x), c_i' + f(x+e_i)\} + x_i(c_i + f(x))\\
&+ (N_i - x_i - 1)\min\{c_i + f(x+e_i+e_j), c_i' + f(x+2e_i+e_j)\}\\
&\qquad + (x_i + 1)(c_i + f(x+e_i+e_j))\\
&- (N_i - x_i - 1)\min\{c_i + f(x+e_i), c_i' + f(x+2e_i)\}\\
&\qquad - (x_i + 1)(c_i + f(x+e_i))\\
&- (N_i - x_i)\min\{c_i + f(x+e_j), c_i' + f(x+e_i+e_j)\} - x_i(c_i + f(x+e_j))\\
=&(N_i - x_i - 1)\\
&\times[\underbrace{\min\{c_i + f(x), c_i' + f(x+e_i)\}}_{(1)} - \underbrace{\min\{c_i + f(x+e_j), c_i' + f(x+e_i+e_j)\}}_{(2)}]\\
&- \underbrace{\min\{c_i + f(x+e_i), c_i' + f(x+2e_i)\}}_{(3)}\\
&+ \underbrace{\min\{c_i + f(x+e_i+e_j), c_i' + f(x+2e_i+e_j)\}}_{(4)}
\end{aligned}
$$

$$
\begin{aligned}
&- \underbrace{\min\{c_i + f(x+e_j), c_i' + f(x+e_i+e_j)\}}_{(5)} + \underbrace{\min\{c_i + f(x), c_i' + f(x+e_i)\}}_{(6)} \\
&+ x_i[f(x) + f(x+e_i+e_j) - f(x+e_i) - f(x+e_j)] \\
&+ f(x+e_i+e_j) - f(x+e_i) \\
\geq & 0.
\end{aligned}
$$

The validity of the last inequality follows by letting the actions in (2), (3) and (5) copy the minimisers of (1), (4) and (6) respectively. Case checking shows that the resulting expression can be reduced to a supermodular expression. The assumption that $f \in \mathcal{S}uper^N(i,j)$ yields the result.

*Proof of 4.* Suppose that $f \in \mathcal{DI}^N(i-1))$. Let $x$ be such that $x_i \leq N_i - 1$. Then it holds that

$$
\begin{aligned}
&N_i(T^S_{CA(i)} f(x+e_i) - T^S_{CA(i)} f(x+e_{i-1})) \\
&\quad = (N_i - x_i - 1)\min\{c_i + f(x+e_i), c_i' + f(x+2e_i)\} \\
&\qquad\qquad + (x_i+1)(c_i + f(x+e_i)) \\
&\qquad - (N_i - x_i)\min\{c_i + f(x+e_{i-1}), c_i' + f(x+e_i+e_{i-1})\} \\
&\qquad\qquad - x_i(c_i + f(x+e_{i-1})) \\
&\quad = (N_i - x_i - 1)\Big[\underbrace{\min\{c_i + f(x+e_i), c_i' + f(x+2e_i)\}}_{(1)} \\
&\qquad\qquad - \underbrace{\min\{c_i + f(x+e_{i-1}), c_i' + f(x+e_i+e_{i-1})\}}_{(2)}\Big] \\
&\qquad - \underbrace{\min\{c_i + f(x+e_{i-1}), c_i' + f(x+e_i+e_{i-1})\}}_{(3)} \\
&\qquad + x_i[f(x+e_i) - f(x+e_{i-1})] + (c_i + f(x+e_i)) \\
&\quad \geq 0.
\end{aligned}
$$

To verify this, choose in (2) the same action as (1), in (3) choose reject. Then the inequality follows since $f \in \mathcal{DI}^N(i-1)$. ☐

**Corollary 7.4.4.** *Consider $\mathcal{T}^N_{SCA}$, let $N \in \mathcal{N}$, then*

*1.*

$$
\mathcal{T}^N_{SCA} : \mathcal{I}^N \to \mathcal{I}^N,
$$

2.
$$\mathcal{T}^N_{SCA} : \mathcal{C}x^N \to \mathcal{C}x^N,$$

3.
$$\mathcal{T}^N_{SCA} : \mathcal{S}uper^N \to \mathcal{S}uper^N.$$

*Proof.* This is a direct consequence of Theorem 7.4.3. □

Notice that we have no convergence result for $\mathcal{DI}^N$, since it is not clear how to derive this for $\mathcal{T}^N_{SCA}$.

### 7.4.3 SRT Double Arrivals

**Theorem 7.4.5.** *Let $N \in \mathcal{N}$, then for $i \neq j \in \{1, \ldots, K\}$,*

1.
$$\mathcal{T}^N_{SDA(i,j)} : \mathcal{I}^N(i) \to \mathcal{I}^N(i),$$

2.
$$\mathcal{T}^N_{SDA(i,j)} : \mathcal{C}x^N(i) \cap \mathcal{S}uper^N(i,j) \to \mathcal{C}x^N(i),$$

3.
$$\mathcal{T}^N_{SDA(i,j)} : \mathcal{S}uper^N(i,j) \to \mathcal{S}uper^N(i,j),$$

4. *if $j = i+1 \leq K$ then*
$$\mathcal{T}^N_{SDA(i,j)} : \mathcal{DI}^N(i) \to \mathcal{DI}^N(i).$$

*Proof.* The results are from [16, p.47]. Let $x \in \mathbf{S}^N$, let $i \neq j \in \{1, \ldots, K\}$.

*Proof of 1.* Suppose that $f \in \mathcal{I}^N(i)$, and let $x+e_i \in \mathbf{S}^N$. If $x_i+x_j \leq N_{ij}-1$, then we have

$$\begin{aligned}
&N_{ij}\big(\mathcal{T}^N_{SDA(i,j)}f(x+e_i) - \mathcal{T}^N_{SDA(i,j)}f(x)\big)\\
=&(N_{ij} - x_i - x_j - 1)f(x+2e_i+e_j) + x_j f(x+2e_i) + (x_i+1)f(x+e_i+e_j)\\
&- (N_{ij} - x_i - x_j)f(x+e_i+e_j) - x_j f(x+e_i) - x_i f(x+e_j)\\
=&(N_{ij} - x_i - x_j - 1)[f(x+2e_i+e_j) - f(x+e_i+e_j)] - f(x+e_i+e_j)\\
&+ x_j[f(x+2e_i) - f(x+e_i)]\\
&+ x_i[f(x+e_i+e_j) - f(x+e_j)] + f(x+e_i+e_j)\\
\geq&0.
\end{aligned}$$

The inequality follows from $f \in \mathcal{I}^N(i)$.

Next suppose $x_i + x_j \geq N_{ij}$ (and $x_i \leq N_{ij} - 1, x_j \leq N_{ij}$), then

$$\begin{aligned}
N_{ij}&\big(\mathcal{T}^N_{SDA(i,j)} f(x+e_i) - \mathcal{T}^N_{SDA(i,j)} f(x)\big) \\
&= (N_{ij} - x_i - 1) f(x+2e_i) + (N_{ij} - x_j) f(x+e_i+e_j) \\
&\qquad\qquad + (x_i + x_j - N_{ij} + 1) f(x+e_i) \\
&\quad - (N_{ij} - x_i) f(x+e_i) - (N_{ij} - x_j) f(x+e_j) - (x_i + x_j - N_{ij}) f(x) \\
&= (N_{ij} - x_i - 1)[f(x+2e_i) - f(x+e_i)] - f(x+e_i) \\
&\quad + (N_{ij} - x_j)[f(x+e_i+e_j) - f(x+e_j)] \\
&\quad (x_i + x_j - N_{ij})[f(x+e_i) - f(x)] + f(x+e_i) \\
&\geq 0,
\end{aligned}$$

and the inequality follows from $f \in \mathcal{I}^N(i)$.

*Proof of 2.* Suppose $f \in \mathcal{C}x^N(i) \cap \mathcal{S}uper^N(i,j)$. Assume $x + 2e_i \in \mathbf{S}^N$. If $x_i + x_j \leq N_{ij} - 2$, we have

$$\begin{aligned}
N_{ij}&\big(\mathcal{T}^N_{SDA(i,j)} f(x+2e_i) - 2\mathcal{T}^N_{SDA(i,j)} f(x+e_i) + \mathcal{T}^N_{SDA(i,j)} f(x)\big) \\
=&(N_{ij} - x_i - x_j - 2) f(x+3e_i+e_j) + x_j f(x+3e_i) + (x_i+2) f(x+2e_i+e_j) \\
&- 2(N_{ij} - x_i - x_j - 1) f(x+2e_i+e_j) - 2x_j f(x+2e_i) \\
&\qquad - 2(x_i+1) f(x+e_i+e_j) \\
&+ (N_{ij} - x_i - x_j) f(x+e_i+e_j) + x_j f(x+e_i) + x_i f(x+e_j) \\
=&(N_{ij} - x_i - x_j - 2)[f(x+3e_i+e_j) - 2f(x+2e_i+e_j) + f(x+e_i+e_j)] \\
&+ x_j[f(x+3e_i) - 2f(x+2e_i) + f(x+e_i)] \\
&+ x_i[f(x+2e_i+e_j) - 2f(x+e_i+e_j) + f(x+e_j)] \\
&- 2f(x+2e_i+e_j) + 2f(x+e_i+e_j) \\
&+ 2f(x+2e_i+e_j) - 2f(x+e_i+e_j) \\
\geq& 0.
\end{aligned}$$

The inequality follows from $\mathcal{C}x^N(i)$.

For $x_i + x_j = N_{ij} - 1$ (and $x_i \leq N_{ij} - 2$) we have

$$\begin{aligned}
N_{ij}&\big(\mathcal{T}^N_{SDA(i,j)} f(x+2e_i) - 2\mathcal{T}^N_{SDA(i,j)} f(x+e_i) + \mathcal{T}^N_{SDA(i,j)} f(x)\big) \\
&= (x_j - 1) f(x+3e_i) + (N_{ij} - x_j) f(x+2e_i+e_j) + f(x+2e_i) \\
&\quad - 2x_j f(x+2e_i) - 2(N_{ij} - x_j) f(x+e_i+e_j)
\end{aligned}$$

$$
\begin{aligned}
&+ f(x+e_i+e_j) + x_j f(x+e_i) + (N_{ij}-x_j-1) f(x+e_j)\\
=&(x_j-1)[f(x+3e_i) - 2f(x+2e_i) + f(x+e_i)]\\
&+ (N_{ij}-x_j-1)[f(x+2e_i+e_j) - 2f(x+e_i+e_j) + f(x+e_j)]\\
&+ f(x+2e_i+e_j) - 2f(x+e_i+e_j)\\
&- 2f(x+2e_i) + f(x+e_i)\\
&+ f(x+2e_i) + f(x+e_i+e_j)\\
\geq& f(x+e_i) + f(x+2e_i+e_j) - f(x+2e_i) - f(x+e_i+e_j)\\
\geq& 0.
\end{aligned}
$$

The first inequality comes from $f \in \mathcal{C}x^N(i)$, the second inequality follows from $f \in \mathcal{S}uper^N(i,j)$.

For $x_i + x_j \geq N_{ij}$ (and $x_i \leq N_{ij} - 2, x_j \leq N_{ij}$) we get

$$
\begin{aligned}
&N_{ij}\big(\mathcal{T}^N_{SDA(i,j)} f(x+2e_i) - 2\mathcal{T}^N_{SDA(i,j)} f(x+e_i) + \mathcal{T}^N_{SDA(i,j)} f(x)\big)\\
&\quad=(N_{ij}-x_i-2) f(x+3e_i) + (N_{ij}-x_j) f(x+2e_i+e_j)\\
&\qquad\qquad + (x_i+x_j-N_{ij}+2) f(x+2e_i)\\
&\qquad - 2(N_{ij}-x_i-1) f(x+2e_i) - 2(N_{ij}-x_j) f(x+e_i+e_j)\\
&\qquad\qquad - 2(x_i+x_j-N_{ij}+1) f(x+e_i)\\
&\qquad + (N_{ij}-x_i) f(x+e_i) + (N_{ij}-x_j) f(x+e_j) + (x_i+x_j-N_{ij}) f(x)\\
&\quad=(N_{ij}-x_i-2)[f(x+3e_i) - 2f(x+2e_i) + f(x+e_i)]\\
&\qquad + (N_{ij}-x_j)[f(x+2e_i+e_j) - 2f(x+e_i+e_j) + f(x+e_j)]\\
&\qquad + (x_i+x_j-N_{ij})[f(x+2e_i) - 2f(x+e_i) + f(x)]\\
&\qquad - 2f(x+2e_i) + 2f(x+e_i)\\
&\qquad + 2f(x+2e_i) - 2f(x+e_i)\\
&\quad\geq 0.
\end{aligned}
$$

The inequality follows from $f \in \mathcal{C}x^N(i)$.

*Proof of 3.* Suppose $f \in \mathcal{S}uper^N(i,j)$ and let $x + e_i + e_j \in \mathbf{S}^N$. If $x_i + x_j \leq N_{ij} - 2$, then

$$
\begin{aligned}
&N_{ij}\big(\mathcal{T}^N_{SDA(i,j)} f(x) + \mathcal{T}^N_{SDA(i,j)} f(x+e_i+e_j)\\
&\qquad\qquad - \mathcal{T}^N_{SDA(i,j)} f(x+e_i) - \mathcal{T}^N_{SDA(i,j)} f(x+e_j)\big)\\
=&(N_{ij}-x_i-x_j) f(x+e_i+e_j) + x_j f(x+e_i) + x_i f(x+e_j)\\
&+ (N_{ij}-x_i-x_j-2) f(x+2e_i+2e_j) + (x_j+1) f(x+2e_i+e_j)
\end{aligned}
$$

$$
\begin{aligned}
&\qquad\quad + (x_i+1)f(x+e_i+2e_j)\\
&\quad -(N_{ij}-x_i-x_j-1)f(x+2e_i+e_j) - x_jf(x+2e_i) - (x_i+1)f(x+e_i+e_j)\\
&\quad -(N_{ij}-x_i-x_j-1)f(x+e_i+2e_j) - (x_j+1)f(x+e_i+e_j) - x_if(x+2e_j)\\
=&(N_{ij}-x_i-x_j-2)[f(x+e_i+e_j)+f(x+2e_i+2e_j)-f(x+2e_i+e_j)\\
&\qquad\quad -f(x+e_i+2e_j)]\\
&\quad + x_j[f(x+e_i)+f(x+2e_i+e_j)-f(x+2e_i)-f(x+e_i+e_j)]\\
&\quad + x_i[f(x+e_j)+f(x+e_i+2e_j)-f(x+e_i+e_j)-f(x+2e_j)]\\
&\quad + 2f(x+e_i+e_j)-f(x+2e_i+e_j)-f(x+e_i+2e_j)\\
&\quad + f(x+2e_i+e_j)-f(x+e_i+e_j)+f(x+e_i+2e_j)-f(x+e_i+e_j)\\
\geq&0.
\end{aligned}
$$

The compensation terms cancel, and so the inequality follows from $f \in \mathcal{S}uper^N(i,j)$.

If $x_i + x_j = N_{ij} - 1$, then we get

$$
\begin{aligned}
&N_{ij}\big(\mathcal{T}^N_{SDA(i,j)}f(x)+\mathcal{T}^N_{SDA(i,j)}f(x+e_i+e_j)\\
&\qquad\quad -\mathcal{T}^N_{SDA(i,j)}f(x+e_i)-\mathcal{T}^N_{SDA(i,j)}f(x+e_j)\big)\\
=&f(x+e_i+e_j)+x_jf(x+e_i)+(N_{ij}-x_j-1)f(x+e_j)\\
&\quad x_jf(x+2e_i+e_j)+(N_{ij}-x_j-1)f(x+e_i+2e_j)+f(x+e_i+e_j)\\
&\quad -x_jf(x+2e_i)-(N_{ij}-x_j)f(x+e_i+e_j)\\
&\quad -(x_j+1)f(x+e_i+e_j)+(N_{ij}-x_j-1)f(x+2e_j)\\
=&x_j[f(x+e_i)+f(x+2e_i+e_j)-f(x+2e_i)-f(x+e_i+e_j)]\\
&\quad +(N_{ij}-x_j-1)[f(x+e_j)+f(x+e_i+2e_j)-f(x+e_i+e_j)-f(x+2e_j)]\\
\geq&0,
\end{aligned}
$$

the inequality follows from $f \in \mathcal{S}uper^N(i,j)$.

For $x_i + x_j \geq N_{ij}$ (and $x_i \leq N_{ij}-1, x_j \leq N_{ij}-1$), we have

$$
\begin{aligned}
&N_{ij}\big(\mathcal{T}^N_{SDA(i,j)}f(x)+\mathcal{T}^N_{SDA(i,j)}f(x+e_i+e_j)\\
&\qquad\quad -\mathcal{T}^N_{SDA(i,j)}f(x+e_i)-\mathcal{T}^N_{SDA(i,j)}f(x+e_j)\big)\\
=&(N_{ij}-x_i)f(x+e_i)+(N_{ij}-x_j)f(x+e_j)+(x_i+x_j-N_{ij})f(x)\\
&\quad +(N_{ij}-x_i-1)f(x+2e_i+e_j)+(N_{ij}-x_j-1)f(x+e_i+2e_j)\\
&\qquad\quad +(x_i+x_j-N_{ij}+2)f(x+e_i+e_j)\\
&\quad -(N_{ij}-x_i-1)f(x+2e_i)-(N_{ij}-x_j)f(x+e_i+e_j)\\
&\qquad\quad -(x_i+x_j-N_{ij}+1)f(x+e_i)
\end{aligned}
$$

$$
\begin{aligned}
&- (N_{ij} - x_i)f(x + e_i + e_j) + (N_{ij} - x_j - 1)f(x + 2e_j) \\
&\qquad\qquad - (x_i + x_j - N_{ij} + 1)f(x + e_j) \\
=&(N_{ij} - x_i - 1)[f(x + e_i) + f(x + 2e_i + e_j) - f(x + 2e_i) - f(x + e_i + e_j)] \\
&+ (N_{ij} - x_j - 1)[f(x + e_j) + f(x + e_i + 2e_j) - f(x + e_i + e_j) - f(x + 2e_j)] \\
&+ (x_i + x_j - N_{ij})[f(x) + f(x + e_i + e_j) - f(x + e_i) - f(x + e_j)] \\
&+ f(x + e_i) - f(x + e_i + e_j) \\
&+ f(x + e_j) - f(x + e_i + e_j) \\
&+ 2f(x + e_i + e_j) - f(x + e_i) - f(x + e_j) \\
\geq&0.
\end{aligned}
$$

The inequality follows from $f \in \mathcal{S}uper^N(i, j)$.

*Proof of 4.* Suppose $f \in \mathcal{DI}^N(i)$, and let $x + e_i, x + e_j \in \mathbf{S}^N$. Then for $x_i + x_j \leq N_{ij} - 1$

$$
\begin{aligned}
&N_{ij}\big(\mathcal{T}^N_{SDA(i,j)} f(x + e_j) - \mathcal{T}^N_{SDA(i,j)} f(x + e_i)\big) \\
=&(N_{ij} - x_i - x_j - 1)f(x + e_i + 2e_j) + (x_j + 1)f(x + e_i + e_j) + x_i f(x + 2e_j) \\
&-(N_{ij} - x_i - x_j - 1)f(x + 2e_i + e_j) - x_j f(x + 2e_i) - (x_i + 1)f(x + e_i + e_j) \\
=&(N_{ij} - x_i - x_j - 1)[f(x + e_i + e_j) - f(x + 2e_i + e_j)] \\
&+ x_j[f(x + e_i + e_j) - f(x + 2e_i)] + f(x + e_i + e_j) \\
&+ x_i[f(x + 2e_j) - f(x + e_i + e_j)] - f(x + e_i + e_j) \\
\geq&0.
\end{aligned}
$$

The inequality is direct from $f \in \mathcal{DI}^N(i)$.

For $x_i + x_j \geq N_{ij}$ and $(x_i, x_j \leq N_{ij} - 1)$ the following expression holds

$$
\begin{aligned}
&N_{ij}\big(\mathcal{T}^N_{SDA(i,j)} f(x + e_j) - \mathcal{T}^N_{SDA(i,j)} f(x + e_i)\big) \\
=&(N_{ij} - x_i)f(x + e_i + e_j) + (N_{ij} - x_j - 1)f(x + 2e_j) \\
&\qquad\qquad + (x_i + x_j - N_{ij} + 1)f(x + e_j) \\
&- (N_{ij} - x_i - 1)f(x + 2e_i) - (N_{ij} - x_j)f(x + e_i + e_j) \\
&\qquad\qquad - (x_i + x_j - N_{ij} + 1)f(x + e_i) \\
=&(N_{ij} - x_i - 1)[f(x + e_i + e_j) - f(x + 2e_i)] + f(x + e_i + e_j) \\
&+ (N_{ij} - x_j - 1)[f(x + 2e_j) - f(x + e_i + e_j)] - f(x + e_i + e_j) \\
&+ (x_i + x_j - N_{ij} + 1)[f(x + e_j) - f(x + e_i)] \\
\geq&0,
\end{aligned}
$$

again $f \in \mathcal{DI}^N(i)$ yields the inequality. $\square$

**Corollary 7.4.6.** *Let $N \in \mathcal{N}$, then for $i \neq j \in \{1, \ldots, K\}$*

1. $$\mathcal{T}^N_{SDA(i,j)} : \mathcal{I}^N \to \mathcal{I}^N,$$

2. $$\mathcal{T}^N_{SDA(i,j)} : \mathcal{C}x^N \cap \mathcal{S}uper^N \to \mathcal{C}x^N,$$

3. $$\mathcal{T}^N_{SDA(i,j)} : \mathcal{S}uper^N \to \mathcal{S}uper^N.$$

*Proof.* This is an immediate consequence of Theorem 7.4.5. $\square$

For property $\mathcal{DI}^N$ there is no direct result with respect to $\mathcal{T}^N_{SDA(i,j)}$.

### 7.4.4 SRT Transfers

**Theorem 7.4.7.** *Let $i \in \{1, \ldots, K-1\}$. Let $N \in \mathcal{N}$, then*

1. $$\mathcal{T}^N_{ST(i)} : \mathcal{I}^N(i) \cap \mathcal{I}^N(i+1) \to \mathcal{I}^N(i) \cap \mathcal{I}^N(i+1),$$

2. $$\mathcal{T}^N_{ST(i)} : \mathcal{C}x^N(i) \cap \mathcal{S}uper^N(i,i+1) \cap \mathcal{UI}^N(i) \overset{(a)}{\to} \mathcal{C}x^N(i);$$
$$\mathcal{T}^N_{ST(i)} : \mathcal{C}x^N(i+1)\mathcal{S}uper^N(i,i+1) \overset{(b)}{\to} \mathcal{C}x^N(i+1),$$

3. $$\mathcal{T}^N_{ST(i)} : \mathcal{S}uper^N(i,i+1) \cap \mathcal{C}x^N(i) \cap \mathcal{UI}^N(i) \to \mathcal{S}uper^N(i,i+1),$$

4. $$\mathcal{T}^N_{ST(i)} : \mathcal{DI}^N(i) \to \mathcal{DI}^N(i),\ \mathcal{UI}^N(i) \to \mathcal{UI}^N(i).$$

*Proof.* The proofs are based on [16, p.39]. Suppose that $x \in \mathbf{S}^N$.

*Proof of 1.* Assume that $f \in \mathcal{I}^N(i) \cap \mathcal{I}^N(i+1)$. First, we prove that $\mathcal{T}^N_{ST(i)} f \in \mathcal{I}^N(i)$. Because the truncation is in the $(i+1)$-th component, the only non-trivial case is when $x_i = 0$. Then we have

$$\begin{aligned} &N_{i+1}(\mathcal{T}^N_{ST(i)} f(x+e_i) - \mathcal{T}^N_{ST(i)} f(x)) \\ &\quad =(N_{i+1} - x_{i+1}) f(x+e_{i+1}) + x_{i+1} f(x+e_i) \\ &\qquad - N_{i+1} f(x) \\ &\quad =(N_{i+1} - x_{i+1})[f(x+e_{i+1}) - f(x)] \\ &\qquad + x_{i+1}[f(x+e_i) - f(x)] \\ &\quad \geq 0. \end{aligned}$$

The inequality follows from $f \in \mathcal{I}^N(i) \cap \mathcal{I}^N(i+1)$.

Next we prove that $\mathcal{T}^N_{ST(i)} f \in \mathcal{I}^N(i+1)$. For $x_i > 0, x_{i+1} \leq N_{i+1} - 1$, we have

$$\begin{aligned} &N_{i+1}(\mathcal{T}^N_{ST(i)} f(x+e_{i+1}) - \mathcal{T}^N_{ST(i)} f(x)) \\ &=(N_{i+1} - x_{i+1} - 1) f(x - e_i + 2e_{i+1}) + (x_{i+1}+1) f(x+e_{i+1}) \\ &\quad - (N_{i+1} - x_{i+1}) f(x - e_i + e_{i+1}) - x_{i+1} f(x) \\ &=(N_{i+1} - x_{i+1} - 1)[f(x - e_i + 2e_{i+1}) - f(x - e_i + e_{i+1})] - f(x - e_i + e_{i+1}) \\ &\quad + x_{i+1}[f(x+e_{i+1}) - f(x)] + f(x+e_{i+1}) \\ &\geq 0, \end{aligned}$$

where again the inequality follows from $f \in \mathcal{I}^N(i) \cap \mathcal{I}^N(i+1)$. The propagation for $x_i = 0$ is trivial, this yields the result.

*Proof of 2(a).* Assume $f \in \mathcal{C}x^N(i) \cap \mathcal{S}uper^N(i, i+1) \cap \mathcal{UI}^N(i)$. The proof is trivial for all states except for the boundary states, when $x_i = 0$. For $x_i = 0$, we have

$$\begin{aligned} &N_{i+1}(\mathcal{T}^N_{ST(i)} f(x+2e_i) - 2\mathcal{T}^N_{ST(i)} f(x+e_i) + \mathcal{T}^N_{ST(i)} f(x)) \\ &\quad =(N_{i+1} - x_{i+1}) f(x+e_i+e_{i+1}) + x_{i+1} f(x+2e_i) \\ &\qquad - 2(N_{i+1} - x_{i+1}) f(x+e_{i+1}) - 2x_{i+1} f(x+e_i) \\ &\qquad + N_{i+1} f(x) \\ &\quad =(N_{i+1} - x_{i+1})(f(x+e_i+e_{i+1}) - 2f(x+e_{i+1}) + f(x)) \\ &\qquad + x_{i+1}[f(x+2e_i) - 2f(x+e_i) + f(x)] \\ &\quad \geq (N_{i+1} - x_{i+1})[f(x+e_i+e_{i+1}) - f(x+e_{i+1}) - f(x+e_i) + f(x)] \\ &\quad \geq 0. \end{aligned}$$

The first inequality follows from $f \in \mathcal{C}x^N(i) \cap \mathcal{UI}^N(i)$. The second inequality follows from $f \in \mathcal{S}uper^N(i, i+1)$.

*Proof of 2(b).* Assume $f \in \mathcal{C}x^N(i+1) \cap \mathcal{S}uper^N(i, i+1)$. Then for $x_{i+1} \leq N_{i+1} - 2$, we get

$$
\begin{aligned}
&N_{i+1}(\mathcal{T}^N_{ST(i)} f(x + 2e_{i+1}) - 2\mathcal{T}^N_{ST(i)} f(x + e_{i+1}) + \mathcal{T}^N_{ST(i)} f(x))\\
=&(N_{i+1} - x_{i+1} - 2) f(x - e_i + 3e_{i+1}) + (x_{i+1} + 2) f(x + 2e_{i+1})\\
&- 2(N_{i+1} - x_{i+1} - 1) f(x - e_i + 2e_{i+1}) - 2(x_{i+1} + 1) f(x + e_{i+1})\\
&+ (N_{i+1} - x_{i+1}) f(x - e_i + e_{i+1}) + x_{i+1} f(x)\\
=&(N_{i+1} - x_{i+1} - 2)[f(x - e_i + 3e_{i+1}) - 2f(x - e_i + 2e_{i+1}) + f(x - e_i + e_{i+1})]\\
&- 2f(x - e_i + 2e_{i+1}) + 2f(x - e_i + e_{i+1})\\
&+ x_{i+1}[f(x + 2e_{i+1}) - 2f(x + e_{i+1}) + f(x)] + 2f(x + 2e_{i+1}) - 2f(x + e_{i+1})\\
\geq& 2[f(x - e_i + e_{i+1}) + f(x + 2e_{i+1}) - f(x - e_i + 2e_{i+1}) - f(x + e_{i+1})]\\
\geq& 0.
\end{aligned}
$$

The first inequality follows from $f \in \mathcal{C}x^N(i)$, the second inequality from $f \in \mathcal{S}uper^N(i, i+1)$.

*Proof of 3.* Suppose $f \in \mathcal{S}uper^N(i, i+1) \cap \mathcal{C}x^N(i) \cap \mathcal{UI}^N(i)$. Then for $x_i > 0,\ x_{i+1} \leq N_{i+1} - 1$, we obtain the following inequality

$$
\begin{aligned}
&N_{i+1}(\mathcal{T}^N_{ST(i)} f(x) + \mathcal{T}^N_{ST(i)} f(x + e_i + e_{i+1})\\
&\qquad - \mathcal{T}^N_{ST(i)} f(x + e_i) - \mathcal{T}^N_{ST(i)} f(x + e_{i+1}))\\
&= (N_{i+1} - x_{i+1})\, f(x - e_i + e_{i+1}) + x_{i+1} f(x)\\
&\quad + (N_{i+1} - x_{i+1} - 1)\, f(x + 2e_{i+1}) + (x_{i+1} + 1) f(x + e_i + e_{i+1})\\
&\quad - (N_{i+1} - x_{i+1})\, f(x + e_{i+1}) - x_{i+1} f(x + e_i)\\
&\quad - (N_{i+1} - x_{i+1} - 1)\, f(x - e_i + 2e_{i+1}) - (x_{i+1} + 1) f(x + e_{i+1})\\
&= (N_{i+1} - x_{i+1} - 1)\,[f(x - e_i + e_{i+1}) + f(x + 2e_{i+1})\\
&\qquad - f(x + e_{i+1}) - f(x - e_i + 2e_{i+1})]\\
&\quad + x_{i+1}[f(x) + f(x + e_i + e_{i+1}) - f(x + e_i) - f(x + e_{i+1})]\\
&\quad + f(x - e_i + e_{i+1}) - 2f(x + e_{i+1}) + f(x + e_i + e_{i+1})\\
&\geq 0.
\end{aligned}
$$

The inequality follows from the assumption that $f \in \mathcal{S}uper^N(i, i+1) \cap \mathcal{C}x^N(i)$.

Now if $x_i = 0$, then the following inequality holds

$$
\begin{aligned}
N_{i+1}(&\mathcal{T}^N_{ST(i)} f(x) + \mathcal{T}^N_{ST(i)} f(x + e_i + e_{i+1}) \\
&\qquad - \mathcal{T}^N_{ST(i)} f(x + e_i) - \mathcal{T}^N_{ST(i)} f(x + e_{i+1})) \\
=&N_{i+1} f(x) \\
&+ (N_{i+1} - x_{i+1} - 1)\, f(x + 2e_{i+1}) + (x_{i+1} + 1) f(x + e_i + e_{i+1}) \\
&- (N_{i+1} - x_{i+1})\, f(x + e_{i+1}) - x_{i+1} f(x + e_i) \\
&- N_{i+1} f(x + e_{i+1}) \\
=&(N_{i+1} - x_{i+1})\, [f(x) + f(x + 2e_{i+1}) - 2f(x + e_{i+1})] \\
&+ x_{i+1}[f(x) + f(x + e_i + e_{i+1}) - f(x + e_i) - f(x + e_{i+1})] \\
&+ f(x - e_i + e_{i+1}) - f(x + 2e_{i+1}) \\
\geq&0.
\end{aligned}
$$

The inequality follows from $f \in \mathcal{S}uper^N(i, i+1) \cap \mathcal{C}x^N(i) \cap \mathcal{UI}^N(i)$.

*Proof of 4.* We only prove $\mathcal{T}^N_{ST(i)} : \mathcal{UI}^N(i) \to \mathcal{UI}^N(i)$. The other propagation follows by replacing '$\geq$' by '$\leq$'. Suppose that $f \in \mathcal{UI}^N(i)$. Then for $x_i > 0,\ x_{i+1} \leq N_{i+1} - 1$ we have

$$
\begin{aligned}
N_{i+1}(&\mathcal{T}^N_{ST(i)} f(x + e_i) - \mathcal{T}^N_{ST(i)} f(x + e_{i+1})) \\
=&(N_{i+1} - x_{i+1}) f(x + e_{i+1}) + x_{i+1} f(x + e_i) \\
&- (N_{i+1} - x_{i+1} - 1) f(x - e_i + 2e_{i+1}) - (x_{i+1} + 1) f(x + e_{i+1}) \\
=&(N_{i+1} - x_{i+1} - 1)[f(x + e_{i+1}) - f(x - e_i + 2e_{i+1})] - f(x + e_{i+1}) \\
&+ x_{i+1}[f(x + e_{i+1}) - f(x - e_i)] + f(x + e_{i+1}) \\
\geq&0.
\end{aligned}
$$

The inequality is valid, since $f \in \mathcal{UI}^N(i)$. For $x_i = 0$ we obtain

$$
\begin{aligned}
N_{i+1}(&\mathcal{T}^N_{ST(i)} f(x + e_i) - \mathcal{T}^N_{ST(i)} f(x + e_{i+1})) \\
=&(N_{i+1} - x_{i+1}) f(x + e_{i+1}) + x_{i+1} f(x + e_i) \\
&- N_{i+1} f(x + e_{i+1}) \\
=&x_{i+1}[f(x + e_{i+1}) - f(x - e_i)] \\
\geq&0.
\end{aligned}
$$

The inequality follows from $f \in \mathcal{UI}^N(i)$. □

**Corollary 7.4.8.** *Let $N \in \mathcal{N}$, then*

1. $$\mathcal{T}_{ST}^{N} : \mathcal{I}^N \to \mathcal{I}^N,$$

2. $$\mathcal{T}_{ST}^{N} : \mathcal{C}x^N \cap \mathcal{S}uper^N \cap \mathcal{UI}^N \to \mathcal{C}x^N,$$

3. $$\mathcal{T}_{ST}^{N} : \mathcal{DI}^N \to \mathcal{DI}^N,\ \mathcal{UI}^N \to \mathcal{UI}^N.$$

*Proof.* This is direct from Theorem 7.4.7. □

The propagation of $\mathcal{S}uper^N$ through $\mathcal{T}_{ST}^{N}$ has not been investigated so far.

### 7.4.5 SRT Increasing Transfers

**Theorem 7.4.9.** *Let $i \in \{1, \ldots, K-1\}$. Let $N \in \mathcal{N}$, then*

1. $$\mathcal{T}_{SIT(i)}^{N} : \mathcal{I}^N(i) \cap \mathcal{I}^N(i+1) \to \mathcal{I}^N(i) \cap \mathcal{I}^N(i+1),$$

2. $$\mathcal{T}_{SIT(i)}^{N} : \mathcal{C}x^N(i) \cap \mathcal{S}uper^N(i, i+1) \to \mathcal{C}x^N(i),$$

3. $$\mathcal{T}_{SIT(i)}^{N} : \mathcal{S}uper^N(i, i+1) \cap \mathcal{C}x^N(i) \cap \mathcal{C}x^N(i+1) \to \mathcal{S}uper^N(i, i+1),$$

4. $$\mathcal{T}_{SIT(i)}^{N} : \mathcal{DI}^N(i) \to \mathcal{DI}^N(i),\ \mathcal{UI}^N(i) \to \mathcal{UI}^N(i).$$

*Proof.* The proofs are from [16, p.43]. Let $x \in \mathbf{S}^N$.

*Proof of 1.* Suppose that $f \in \mathcal{I}^N(i) \cap \mathcal{I}^N(i+1)$. By symmetry it is sufficient to prove $\mathcal{T}^N_{SIT(i)} f \in \mathcal{I}^N(i)$. For $x_i \le N_i - 1$, we have

$$\begin{aligned}
&N_i N_{i+1}\big(\mathcal{T}^N_{SIT(i)} f(x+e_i) - \mathcal{T}^N_{SIT(i)} f(x)\big)\\
&\quad= (N_{i+1} - x_{i+1})(x_i + 1) f(x + e_{i+1}) + (N_i - x_i - 1) x_{i+1} f(x + 2e_i - e_{i+1})\\
&\qquad + \big(N_i N_{i+1} - (N_{i+1} - x_{i+1})(x_i + 1) - (N_i - x_i - 1) x_{i+1}\big) f(x + e_i)\\
&\qquad - (N_{i+1} - x_{i+1}) x_i f(x - e_i + e_{i+1}) - (N_i - x_i) x_{i+1} f(x + e_i - e_{i+1})\\
&\qquad - \big(N_i N_{i+1} - (N_{i+1} - x_{i+1}) x_i - (N_i - x_i) x_{i+1}\big) f(x)\\
&\quad= (N_{i+1} - x_{i+1}) x_i [f(x + e_{i+1}) - f(x - e_i + e_{i+1})]\\
&\qquad + (N_i - x_i - 1) x_{i+1} [f(x + 2e_i - e_{i+1}) - f(x + e_i - e_{i+1})]\\
&\qquad + \big(N_i N_{i+1} - (N_{i+1} - x_{i+1})(x_i + 1) - (N_i - x_i) x_{i+1}\big)[f(x + e_i) - f(x)]\\
&\qquad + (N_{i+1} - x_{i+1})[f(x + e_{i+1}) - f(x)]\\
&\qquad + x_{i+1}[f(x + e_i) - f(x + e_i - e_{i+1})]\\
&\quad\ge 0.
\end{aligned}$$

The inequality follows from $f \in \mathcal{I}^N(i) \cap \mathcal{I}^N(i+1)$.

*Proof of 2.* Suppose $f \in \mathcal{C}x^N(i) \cap \mathcal{S}uper^N(i, i+1)$. For $x_i \le N_i - 2$ we have

$$\begin{aligned}
&N_i N_{i+1}\big(\mathcal{T}^N_{SIT(i)} f(x+2e_i) - 2\mathcal{T}^N_{SIT(i)} f(x+e_i) + \mathcal{T}^N_{SIT(i)} f(x)\big)\\
&= (N_{i+1} - x_{i+1})(x_i + 2) f(x + e_i + e_{i+1}) + (N_i - x_i - 2) x_{i+1} f(x + 3e_i - e_{i+1})\\
&\quad + \big(N_i N_{i+1} - (N_{i+1} - x_{i+1})(x_i + 2) - (N_i - x_i - 2) x_{i+1}\big) f(x + 2e_i)\\
&\quad - 2(N_{i+1} - x_{i+1})(x_i + 1) f(x + e_{i+1}) - 2(N_i - x_i - 1) x_{i+1} f(x + 2e_i - e_{i+1})\\
&\quad - 2\big(N_i N_{i+1} - (N_{i+1} - x_{i+1})(x_i + 1) - (N_i - x_i - 1) x_{i+1}\big) f(x + e_i)\\
&\quad + (N_{i+1} - x_{i+1}) x_i f(x - e_i + e_{i+1}) + (N_i - x_i) x_{i+1} f(x + e_i - e_{i+1})\\
&\quad + \big(N_i N_{i+1} - (N_{i+1} - x_{i+1}) x_i - (N_i - x_i) x_{i+1}\big) f(x)\\
&= (N_{i+1} - x_{i+1}) x_i [f(x + e_i + e_{i+1}) - 2f(x + e_{i+1}) + f(x - e_i + e_{i+1})]\\
&\quad + (N_i - x_i - 2) x_{i+1} [f(x + 3e_i - e_{i+1}) - 2f(x + 2e_i - e_{i+1})\\
&\qquad\qquad + f(x + e_i - e_{i+1})]\\
&\quad + \big(N_i N_{i+1} - (N_{i+1} - x_{i+1})(x_i + 2) - (N_i - x_i) x_{i+1}\big)\\
&\qquad\qquad \times [f(x + 2e_i) - 2f(x + e_i) + f(x)]\\
&\quad + 2(N_{i+1} - x_{i+1})[f(x) + f(x + e_i + e_{i+1}) - f(x + e_i) - f(x + e_{i+1})]\\
&\quad + 2x_{i+1}[f(x + 2e_i) + f(x + e_i - e_{i+1}) - f(x + 2e_i - e_{i+1}) - f(x + e_i)]\\
&\ge 0.
\end{aligned}$$

The inequality follows from $f \in \mathcal{C}x^N(i) \cap \mathcal{S}uper^N(i, i+1)$.

*Proof of 3.* Suppose that $f \in \mathcal{S}uper^N(i, i+1) \cap \mathcal{C}x^N(i) \cap \mathcal{C}x^N(i+1)$. For $x_i \le N_i - 1,\ x_{i+1} \le N_{i+1} - 1$, we get

$$
\begin{aligned}
&N_i N_{i+1} \big( \mathcal{T}^N_{SIT(i)} f(x) + \mathcal{T}^N_{SIT(i)} f(x + e_i + e_{i+1}) \\
&\qquad\qquad - \mathcal{T}^N_{SIT(i)} f(x + e_i) - \mathcal{T}^N_{SIT(i)} f(x + e_{i+1}) \big) \\
=& x_i \left(N_{i+1} - x_{i+1}\right) f(x - e_i + e_{i+1}) + x_{i+1} \left(N_i - x_i\right) f(x + e_i - e_{i+1}) \\
&+ \big( N_i N_{i+1} - x_i \left(N_{i+1} - x_{i+1}\right) - x_{i+1} \left(N_i - x_i\right) \big) f(x) \\
&+ \left(x_i + 1\right) \left(N_{i+1} - x_{i+1} - 1\right) f(x + 2e_{i+1}) \\
&\qquad\qquad + \left(x_{i+1} + 1\right) \left(N_i - x_i - 1\right) f(x + 2e_i) \\
&+ \big( N_i N_{i+1} - \left(x_i + 1\right) \left(N_{i+1} - x_{i+1} - 1\right) \\
&\qquad\qquad - \left(x_{i+1} + 1\right) \left(N_i - x_i - 1\right) \big) f(x + e_i + e_{i+1}) \\
&- \left(x_i + 1\right) \left(N_{i+1} - x_{i+1}\right) f(x + e_{i+1}) - x_{i+1} \left(N_i - x_i - 1\right) f(x + 2e_i - e_{i+1}) \\
&- \big( N_i N_{i+1} - \left(x_i + 1\right) \left(N_{i+1} - x_{i+1}\right) - x_{i+1} \left(N_i - x_i - 1\right) \big) f(x + e_i) \\
&- x_i \left(N_{i+1} - x_{i+1} - 1\right) f(x - e_i + 2e_{i+1}) - \left(x_{i+1} + 1\right) \left(N_i - x_i\right) f(x + e_i) \\
&- \big( N_i N_{i+1} - x_i \left(N_{i+1} - x_{i+1} - 1\right) - \left(x_{i+1} + 1\right) \left(N_i - x_i\right) \big) f(x + e_{i+1}) \\
=& x_i \left(N_{i+1} - x_{i+1} - 1\right) [f(x - e_i + e_{i+1}) + f(x + 2e_{i+1}) \\
&\qquad\qquad - f(x + e_{i+1}) - f(x - e_i + 2e_{i+1})] \\
&+ x_{i+1} \left(N_i - x_i - 1\right) [f(x + e_i - e_{i+1}) + f(x + 2e_i) \\
&\qquad\qquad - f(x + 2e_i - e_{i+1}) - f(x + e_i)] \\
&+ \big( N_i N_{i+1} - \left(x_i + 1\right) \left(N_{i+1} - x_{i+1}\right) - \left(x_{i+1} + 1\right) \left(N_i - x_i\right) \big) \\
&\qquad\qquad \times [f(x) + f(x + e_i + e_{i+1}) - f(x + e_i) - f(x + e_{i+1})] \\
&+ x_i f(x - e_i + e_{i+1}) + (N_{i+1} - x_{i+1} - 1) f(x + 2e_{i+1}) \\
&+ x_{i+1} f(x + e_i - e_{i+1}) + (N_i - x_i - 1) f(x + 2e_i) \\
&+ (N_{i+1} - x_{i+1}) f(x) + (x_i + 1) f(x + e_i + e_{i+1}) \\
&+ (N_i - x_i) f(x) + (x_{i+1} + 1) f(x + e_i + e_{i+1}) \\
&\qquad\qquad - 2(N_{i+1} + x_i - x_{i+1}) f(x + e_{i+1}) \\
&\qquad\qquad - 2(N_{i+1} - x_i + x_{i+1}) f(x + e_i) \\
\ge& x_{i+1} [f(x - e_i + e_{i+1}) - 2f(x + e_{i+1}) + f(x + e_i + e_{i+1})] \\
&+ x_i [f(x + e_i - e_{i+1}) - 2f(x + e_i) + f(x + e_i + e_{i+1})] \\
&+ (N_i - x_i - 1) [f(x) - 2f(x + e_i) + f(x + 2e_i)]
\end{aligned}
$$

$$
\begin{aligned}
&+ (N_{i+1} - x_{i+1} - 1)[f(x) - 2f(x + e_{i+1}) + f(x + 2e_{i+1})] \\
&+ 2[f(x) + f(x + e_i + e_{i+1}) - f(x + e_i) - f(x + e_{i+1})] \\
\geq& 0.
\end{aligned}
$$

The first inequality follows from $f \in \mathcal{S}uper^N(i, i+1)$, the second inequality follows from $f \in \mathcal{S}uper^N(i, i+1) \cap \mathcal{C}x^N(i) \cap \mathcal{C}x^N(i+1)$.

*Proof of 4.* We only prove $\mathcal{T}^N_{SIT(i)} : \mathcal{UI}^N(i) \to \mathcal{UI}^N(i)$, the other result follows by symmetry. Suppose that $f \in \mathcal{UI}^N(i)$. Then for $x_i \leq N_i - 1$, $x_{i+1} \leq N_{i+1} - 1$, we have

$$
\begin{aligned}
&N_i N_{i+1}\big(\mathcal{T}^N_{SIT(i)} f(x + e_i) - \mathcal{T}^N_{SIT(i)} f(x + e_{i+1})\big) \\
&=(N_{i+1} - x_{i+1})(x_i + 1)f(x + e_{i+1}) + (N_i - x_i - 1)x_{i+1} f(x + 2e_i - e_{i+1}) \\
&\quad + \big(N_i N_{i+1} - (N_{i+1} - x_{i+1})(x_i + 1) - (N_i - x_i - 1)x_{i+1}\big) f(x + e_i) \\
&\quad - (N_{i+1} - x_{i+1} - 1)x_i f(x - e_i + 2e_{i+1}) - (N_i - x_i)(x_{i+1} + 1) f(x + e_i) \\
&\quad - \big(N_i N_{i+1} - (N_{i+1} - x_{i+1} - 1)x_i - (N_i - x_i)(x_{i+1} + 1)\big) f(x + e_{i+1}) \\
&=(N_{i+1} - x_{i+1} - 1)x_i[f(x + e_{i+1}) - f(x - e_i + 2e_{i+1})] \\
&\quad + (N_i - x_i - 1)x_{i+1}[f(x + 2e_i - e_{i+1}) - f(x + e_i)] \\
&\quad + \big(N_i N_{i+1} - (N_{i+1} - x_{i+1})(x_i + 1) \\
&\quad - (N_i - x_i)(x_{i+1} + 1)\big)[f(x + e_i) - f(x + e_{i+1})] \\
&\quad + (N_{i+1} - x_{i+1})[f(x + e_{i+1}) - f(x)] \\
&\quad + x_{i+1}[f(x + e_i) - f(x + e_i - e_{i+1})] \\
&\quad + (N_{i+1} - x_{i+1} - 1)f(x + e_{i+1}) + (x_i + 1)f(x + e_{i+1}) \\
&\quad - (N_{i+1} - x_{i+1} - 1)f(x + e_{i+1}) - (x_i + 1)f(x + e_{i+1}) \\
&\quad + (N_i - x_i - 1)f(x + e_i) + (x_{i+1} + 1)f(x + e_i) \\
&\quad - (N_i - x_i - 1)f(x + e_i) - (x_{i+1} + 1)f(x + e_i) \\
&\geq 0.
\end{aligned}
$$

All compensation terms cancel. The inequality follows since $f \in \mathcal{DI}^N(i)$, thus proving the propagation result. □

**Corollary 7.4.10.** *Let $i \in \{1, \ldots, K-1\}$. Let $N \in \mathcal{N}$, then*

*1.*

$$\mathcal{T}^N_{SIT(i)} : \mathcal{I}^N \to \mathcal{I}^N,$$

*2.*
$$\mathcal{T}^N_{SIT(i)} : \mathcal{C}x^N \cap \mathcal{S}uper^N \to \mathcal{C}x^N,$$

*3.*
$$\mathcal{T}^N_{SIT(i)} : \mathcal{S}uper^N \cap \mathcal{C}x^N \to \mathcal{S}uper^N,$$

*4.*
$$\mathcal{T}^N_{SIT(i)} : \mathcal{DI}^N \to \mathcal{DI}^N,\ \mathcal{UI}^N \to \mathcal{UI}^N.$$

*Proof.* This is a direct consequence of Theorem 7.4.9. □

### 7.4.6 Increasing Departures

For this operator the properties are relevant for $\mathbf{S}$ as well as for $\mathbf{S}^N$.

**Theorem 7.4.11.** *Let $i \neq j \in \{1, \ldots, K\}$. Let $N \in \mathcal{N}$, then*

*1.*
$$\mathcal{T}^N_{ID(i)} : \mathcal{I}^N(i) \overset{(a)}{\to} \mathcal{I}^N(i),\ \mathcal{I}(i) \overset{(b)}{\to} \mathcal{I}(i),$$

*2.*
$$\mathcal{T}^N_{ID(i)} : \mathcal{C}x^N(i) \overset{(a)}{\to} \mathcal{C}x^N(i),\ \mathcal{C}x(i) \cap \mathcal{I}(i) \overset{(b)}{\to} \mathcal{C}x(i),$$

*3.*
$$\mathcal{T}^N_{ID(i)} : \mathcal{S}uper^N(i,j) \overset{(a)}{\to} \mathcal{S}uper^N(i,j),\ \mathcal{S}uper(i,j) \overset{(b)}{\to} \mathcal{S}uper(i,j).$$

*Proof.* Results 1(a), 2(a) and 3(a) are from Proposition 5.3.2.

For the rest of the proof, let $i \neq j \in \{1, \ldots, K\}$, and let $x \in \mathbf{S}$.

*Proof of 1(b).* Let $f \in \mathcal{I}(i)$. If $x_i \leq N_i - 1$, then the proof of 1(a) applies. If $x_i \geq N_i$ then $\mathcal{T}^N_{ID(i)} f(x) = f(x - e_i)$ and the propagation is trivial.

*Proof of 2(b).* Let $f \in \mathcal{C}x(i) \cap \mathcal{I}(i)$, if $x_i \leq N_i - 2$, then the proof of 2(a) applies. If $x_i = N_i - 1$, then

$$\begin{aligned}
&N\big(\mathcal{T}^N_{ID(i)} f(x + 2e_i) - 2\mathcal{T}^N_{ID(i)} f(x + e_i) + \mathcal{T}^N_{ID(i)} f(x)\big)\\
&= N[f(x + e_i) - 2f(x) + f(x - e_i)]\\
&\quad + [f(x) - f(x - e_i)]\\
&\geq 0,
\end{aligned}$$

where the inequality follows from $f \in \mathcal{C}x(i) \cap \mathcal{I}(i)$. If $x_i \geq N_i$ then $\mathcal{T}^N_{ID(i)} f(x) = f(x - e_i)$ and the propagation is trivial.

*Proof of 3(b).* Let $f \in \mathcal{S}uper(i,j)$, if $x_i \leq N_i - 1$, then the proof of 3(a) applies. If $x_i \geq N_i$ then $\mathcal{T}^N_{ID(i)} f(x) = f(x - e_i)$ and the propagation is trivial. □

**Corollary 7.4.12.** *Consider Let $N \in \mathcal{N}$. Then*

*1.*
$$\mathcal{T}^N_{ID} : \mathcal{I}^N \to \mathcal{I}^N,\ \mathcal{I} \to \mathcal{I},$$

*2.*
$$\mathcal{T}^N_{ID} : \mathcal{C}x^N \to \mathcal{C}x^N,\ \mathcal{C}x \cap \mathcal{I} \to \mathcal{C}x,$$

*3.*
$$\mathcal{T}^N_{ID} : \mathcal{S}uper^N \to \mathcal{S}uper^N,\ \mathcal{S}uper \to \mathcal{S}uper.$$

*Proof.* This follows directly from Theorem 7.4.11. □

## 7.4.7 Increasing Idle/Off

**Theorem 7.4.13.** *Let $N \in \mathcal{N}$, for $i < K$*

*1.*
$$\mathcal{T}^N_{I/O} : \mathcal{S}uper(i, i+1) \cap \mathcal{S}uper\mathcal{C}(i, i+1) \to \mathcal{S}uper(i, i+1),$$

*2.*
$$\mathcal{T}^N_{I/O} : \mathcal{S}uper(i, i+1) \cap \mathcal{S}uper\mathcal{C}(i, i+1) \to \mathcal{S}uper\mathcal{C}(i, i+1).$$

*Proof.* 1. and 2. are equal to Assertions 4 and 5 from Lemma 4.5.1. □

# Bibliography


[1] I.J.B.F. Adan, V.G. Kulkarni, and A.C.C. van Wijk. Optimal control of a server farm. *INFOR*, 51(4):241–252, 2013.

[2] E. Altman, A. Hordijk, and F.M. Spieksma. Contraction conditions for average and $\alpha$-discount optimality in countable Markov games with unbounded rewards. *Math. Operat. Res.*, 22:588–619, 1997.

[3] W.J. Anderson. *Continuous-Time Markov Chains*. Springer-Verlag, 1991.

[4] T.M. Apostol. *Mathematical analysis*. Addison Wesley Publishing Company, 1974.

[5] R. Atar, C. Giat, and N. Shimkin. The $c\mu/\theta$ rule for many-server queues with abandonment. *Operations Research*, 58(5):1427–1439, 2010.

[6] Y. Aviv and A. Federgruen. The value iteration method for countable state Markov decision processes. *Operat. Res. Let.*, 24:223–234, 1999.

[7] U. Ayesta, P. Jacko, and V. Novak. A nearly-optimal index rule for scheduling of users with abandonment. In *INFOCOM, 2011 Proceedings IEEE*, pages 2849–2857. IEEE, 2011.

[8] J.S. Baras, D.J. Ma, and A.M. Makowski. K competing queues with geometric service requirements and linear costs: The [mu] c-rule is always optimal. *Systems & control letters*, 6(3):173–180, 1985.

[9] R. Bellman. A Markovian decision process. Technical report, DTIC Document, 1957.

[10] S. Bhulai, H. Blok, and F.M. Spieksma. Optimal control for K-competing queues with abandonments. In preparation, 2016.

[11] S. Bhulai, A.C. Brooms, and F.M. Spieksma. On structural properties of the value function for an unbounded jump Markov process with an application to a processor sharing retrial queue. *Queueing Syst.*, 76(4):425–446, 2014.

[12] S. Bhulai and G. Koole. On the structure of value functions for threshold policies in queueing models. *J. Appl. Prob.*, 40(3):613–622, 2003.

[13] S. Bhulai and F.M. Spieksma. On the uniqueness of solutions to the Poisson equations for average cost Markov chains with unbounded cost functions. *Math. Meth. Operat. Res.*, 58(2):221–236, 2003.

[14] P. Billingsley. *Convergence of Probability Measures.* Wiley Series in Probability and Statistics. J. Wiley & Sons, New York, 2d edition edition, 1999.

[15] D. Blackwell. Discrete dynamic programming. *Ann. Math. Stat.*, 33(2):719–726, 1962.

[16] H. Blok. Markov decision processes with unbounded transition rates: structural properties of the relative value function. Master's thesis, Utrecht University, 2011.

[17] H. Blok, R. Kappetein, and F.M. Spieksma. Power control of a server farm: switching curve optimality. In preparation, 2016.

[18] H. Blok and F.M. Spieksma. Countable state Markov decision processes with unbounded jump rates and discounted cost: optimality and approximations. *Adv. Appl. Probab.*, 40, 2015.

[19] H. Blok and F.M. Spieksma. Roadmap to structures of optimal policies in Markov Decision Processes with unbounded jumps. In R.J. Boucherie and N.M. van Dijk, editors, *MDPs in Practice.* 2016.

[20] V.S. Borkar. Control of Markov chains with long-run average cost criterion: the dynamic programming equations. *SIAM Journal on Control and Optimization*, 27(3):642–657, 1989.

[21] V.S. Borkar. *Topics in controlled Markov chains.* Longman Sc & Tech, 1991.

[22] C. Buyukkoc, P. Varaiya, and J. Walrand. The c$\mu$ rule revisited. *Adv. Appl. Probab.*, 17(1):237–238, 1985.

[23] R. Dekker and A. Hordijk. Average, sensitive and Blackwell optimal policies in denumerable Markov decision chains with unbounded rewards. *Math. Operat. Res.*, 13:395–421, 1988.

[24] R. Dekker and A. Hordijk. Recurrence conditions for average and Blackwell optimality in denumerable Markov decision chains. *Math. Operat. Res.*, 17:271–289, 1992.

[25] R. Dekker, A. Hordijk, and F.M. Spieksma. On the relation between recurrence and ergodicity properties in denumerable Markov decision chains. *Math. Operat. Res.*, 19:539–559, 1994.

[26] C. Derman. *Finite state Markovian decision processes.* Academic Press, New York, 1970.

[27] D.G. Down, G. Koole, and M.E. Lewis. Dynamic control of a single-server system with abandonments. *Queueing Syst.*, 67(1):63–90, 2011.

[28] W. Ellens, M. Mandjes, H. van den Berg, D. Worms, and S. Blaszczuk. Inference of short time-scale fluctuainference of short time-scale fluctuainference of short time-scale fluctuainference of short time-scale fluctuations from periodic system-state measurements. *Perf. Eval.*, 29(C):27–46, 2015.

[29] A. Federgruen. On n-person stochastic games with denumerable state space. *Adv. Appl. Probab.*, 10(2):pp. 452–471, 1978.

[30] E.A. Feinberg. Total reward criteria. In E.A. Feinberg and A. Shwartz, editors, *Handbook of Markov Decision Processes*, volume 40 of *International Series in Operations Research and Management Science*, chapter 5, pages 155–189. Kluwer Academic Publishers, Amsterdam, 2002.

[31] L. Fisher and S.M. Ross. An example in denumerable decision processes. *Ann. Math. Stat.*, 39(2):674–675, 1968.

[32] A. Gandhi, V. Gupta, M. Harchol-Balter, and M.A. Kozuch. Optimality analysis of energy-performance trade-off for server farm management. *Performance Evaluation*, 67(11):1155–1171, 2010.

[33] A. Gandhi, M. Harchol-Balter, and I.J.B.F. Adan. Server farms with setup costs. *Performance Evaluation*, 67(11):1123–1138, 2010.

[34] X. Guo and O. Hernández-Lerma. Continuous-time controlled markov chains with discounted rewards. *Acta Applicandae Mathematica*, 79(3):195–216, 2003.

[35] X. Guo and O. Hernández-Lerma. *Continuous-Time Markov Decision Processes*, volume 62 of *Stochastic Modelling and Applied Probability.* Springer-Verlag, 2009.

[36] X. Guo, O. Hernández-Lerma, and T. Prieto-Rumeau. A survey of recent results on continuous-time markov decision processes. *Top*, 14(2):177–261, 2006.

[37] X. Guo and A. Piunovskiy. Discounted continuous-time markov decision processes with constraints: unbounded transition and loss rates. *Mathematics of Operations Research*, 36(1):105–132, 2011.

[38] X. Guo and W. Zhu. Denumerable-state continuous-time markov decision processes with unbounded transition and reward rates under the discounted criterion. *J. Appl. Probab.*, pages 233–250, 2002.

[39] A. Hordijk. *Dynamic Programming and Markov potential Theory.* Math. Centre Tracts **51**. C.W.I., Amsterdam, 1974.

[40] A. Hordijk. Regenerative Markov decision models. *Math. Program. Study*, pages 49–72, 1976.

[41] A. Hordijk, P.J. Schweitzer, and H.C. Tijms. The asymptotic behaviour of the minimal total expected cost for the denumerable state Markov decision model. *J. Appl. Prob.*, 12(298–305), 1975.

[42] R. Kappetein. Optimal control of a server farm. Master's thesis, Leiden University, 2014.

[43] M.N. Katehakis, L.C. Smit, and F.M. Spieksma. DES and RES processes and their explicit solutions. *Prob. Engin. Inform. Sci.*, 29(2):191–217, 2015.

[44] M. Kitaev. Semi-Markov and jump Markov controlled models: average cost criterion. *Theory Prob. Appl.*, 30:272–288, 1986.

[45] G. Koole. *Monotonicity in Markov reward and decision chains: Theory and Applications*, volume 1. Now Publishers Inc., 2007.

[46] M. Larranaga, U. Ayesta, and I.M. Verloop. Dynamic fluid-based scheduling in a multi-class abandonment queue. *Performance Evaluation*, 70(10):841–858, 2013.

[47] S.A. Lippman. On dynamic programming with unbounded rewards. *Management Science*, 21:1225–1233, 1975.

[48] V.J. Maccio and D.G. Down. On optimal policies for energy-aware servers. *Performance Evaluation*, 2015.

[49] S.P. Meyn and R.L. Tweedie. *Markov Chains and Stochastic Stability.* Springer-Verlag, Berlin, 1993.

[50] J. Munkres. *Topology.* Prenctice Hall, US, 2000.

[51] A. Piunovskiy and Y. Zhang. Discounted continuous-time Markov decision processes with unbounded rates and history dependent policies: the dynamic programming approach. *4OR-Q J. Oper. Res.*, 12:49–75, 2014.

[52] T. Prieto-Rumeau and O. Hernández-Lerma. Discounted continuous-time controlled Markov chains: convergence of control models. *J. Appl. Probab.*, 49(4):1072–1090, 2012.

[53] T. Prieto-Rumeau and O. Hernández-Lerma. *Selected topics on continuous-time controlled Markov chains and Markov games*, volume 5. World Scientific, 2012.

[54] M.L. Puterman. *Markov Decision Processes: Discrete Stochastic Programming.* J. Wiley & Sons, New Jersey, 2d edition edition, 2005.

[55] G.E.H. Reuter. Denumerable Markov processes and the associated contraction semigroups on $l$. *Acta Math.*, 97:1–46, 1957.

[56] S.M. Ross. Non-discounted denumerable Markovian decision models. *Ann. Math. Statist.*, 39(2):412–423, 1968.

[57] S.M. Ross. *Applied Probability Models with Optimization Applications.* Holden Day, San Francisco, 1970.

[58] H.L. Royden. *Real Analysis.* Macmillan Publishing Company, New York, 2d edition, 1988.

[59] W. Rudin. *Principles of Mathematical Analysis.* McGraw-Hill, Inc., 3d edition edition, 1976.

[60] A. Salch, J.-P. Gayon, and P. Lemaire. Optimal static priority rules for stochastic scheduling with impatience. *Operations Research Letters*, 41(1):81–85, 2013.

[61] L.I. Sennott. Average cost optimal stationary policies in infinite state Markov decision processes with unbounded costs. *Operat. Res.*, 37(4):626–633, 1989.

[62] L.I. Sennott. Valute iteration in countable state average cost Markov decision processes with unbounded costs. *Ann. Oper. Res.*, 28:261–272, 1991.

[63] L.I. Sennott. *Stochastic Dynamic Programming and the Control of Queueing Systems.* Wiley Series in Probability and Statistics. Wiley, New York, 1999.

[64] R.F. Serfozo. An equivalence between continuous and discrete time Markov decision processes. *Operat. Res.*, 27(3):616–620, 1979.

[65] F.M. Spieksma. *Geometrically ergodic Markov Chains and the optimal Control of Queues.* PhD thesis, Leiden University, 1990. Available on request from the author.

[66] F.M. Spieksma. Kolmogorov forward equation and explosiveness in countable state Markov processes. *Ann. Operat. Res.*, 2013. DOI: 10.1007/s10479-012-1262-7.

[67] F.M. Spieksma. Countable state Markov processes: non-explosiveness and moment function. *Prob. Engin. Inform. Sci.*, 29(4):623–637, 2015.

[68] F.M. Spieksma. Countable state Markov processes: non-explosiveness and moment function. Accepted for publication in PEIS, 2015.

[69] F.M. Spieksma. Parametrised countable state Markov processes in discrete and continuous time: drift conditions for exponential convergence. In preparation, 2015.

[70] R.E. Strauch. Negative Dynamic Programming. *Ann. Math. Stat.*, 37(4):871–890, 1966.

[71] E.C. Titchmarsh. *The Theory of Functions.* Oxford University Press, 2d edition, 1986.

[72] R.L. Tweedie. Criteria for ergodicity, exponential ergodicity and strong ergodicity of Markov processes. *J. Appl. Prob.*, pages 122–130, 1981.

[73] O.J.G. van der Velde. Optimaal stroomverbruik in een server farm. Bachelor's thesis, Leiden University, 2014.

[74] R.R. Weber and S. Stidham Jr. Optimal control of service rates in networks of queues. *Adv. Appl. Prob.*, 19(1):202–218, 1987.

[75] J. Wessels. Markov programming by successive approximations with respect to weighted supremum norms. *J. Math. An. Appl.*, 58(2):326–335, 1977.

[76] D.V. Widder. *The Laplace Transform.* Princeton University Press, London, 1941.

# Nederlandse Samenvatting

## Markov beslissingsprocessen met onbegrensde sprongintensiteit:

**structuureigenschappen via parametrisatie**

Een Markov beslissingsproces is een stochastisch proces, waarbij de overgangen worden beïnvloed door een serie van gekozen acties. De serie acties wordt bepaald door een strategie die gekozen wordt met het doel de verwachte kosten, horend bij het proces, te minimaliseren. Toepassingen zijn onder meer te vinden in wachtrij problemen. Een intrinsieke eigenschap van Markov beslissingsprocessen is dat er een afweging gemaakt moet worden tussen kosten op korte termijn en het verwachte effect op de lange termijn. Dit zorgt ervoor dat het vinden van een optimale strategie in het algemeen niet eenvoudig is.

In dit proefschrift worden methodes beschreven voor het afleiden van structuureigenschappen van Markov beslissingsprocessen. Als de waardefunctie van een Markov beslissingsproces bepaalde monotonie-eigenschappen bezit, zoals stijgendheid of convexiteit, dan geeft dit vaak informatie over de optimale strategie van het proces. Wat bedoeld wordt met een optimale strategie hangt af van het criterium dat men beschouwt. In dit proefschrift behandelen we vooral de verwachte verdisconteerde kosten en de verwachte gemiddelde kosten.

In ons onderzoek hebben we ons in het bijzonder gericht op continue-tijd Markov beslissingsprocessen met een zogenaamde onbegrensde overgangsintensiteit, omdat hiervoor nog geen systematische manier is om eigenschappen af te leiden. Voor Markov beslissingsprocessen in discrete tijd bestaat er wel een zeer krachtig middel om structuureigenschappen aan te tonen in de vorm van waarde iteratie. Door middel van uniformisatie kan een proces in continue-tijd met een vaste bovengrens op de overgangsintensiteit worden vertaald naar een equivalent discrete-tijd beslissingsproces. Hierdoor komt waarde iteratie ook voor deze processen beschikbaar. Voor Markov beslissingsprocessen met onbegrensde intensiteit is dit niet direct mogelijk met behoud van de gewenste structuureigenschappen.

Het is mogelijk om een Markov beslissingsproces met onbegrensde intensiteiten uniformiseerbaar te maken door middel van een afkapping of afknotting. Het geüniformiseerde afgekapte proces kan dan geanalyseerd worden met waarde iteratie. Dit proces geldt dan als een benadering van het originele proces. In Hoofdstuk 3 leiden we milde drift-voorwaarden af onder welke de benadering convergeert naar het originele proces. Dit wordt gedaan voor het verdisconteerde-kostencriterium.

Het gemiddelde-kosten probleem is lastiger, hierdoor lijkt een gelijksoortige convergentiestelling buiten zicht, zonder al te zware voorwaarden te stellen aan het beslissingsproces. In Hoofdstuk 2 beschrijven we een manier om eigenschappen voor het gemiddelde-kosten probleem aan te tonen door de verdisconteringsfactor naar nul te laten gaan. De eigenschappen volgen als de limiet van het verdisconteerde-kosten probleem.

De twee genoemde limietstellingen stellen ons in staat om modellen met onbegrensde intensiteiten te bestuderen. Via afkappingen kunnen de modellen uniformiseerbaar worden gemaakt. Een complicerende factor is echter dat door randeffecten van de afkapping monotonie-eigenschappen verloren kunnen gaan.

In Hoofdstuk 4 bestuderen we een serverfarm waarbij het doel is kosten te minimaliseren door een overschot aan servers op standby te beperken. Een standaardafknotting van de sprongintensiteit voldoet hier, want deze behoudt de monotonie-eigenschappen. Uit deze eigenschappen kunnen we afleiden dat een drempelstrategie optimaal is. Waarbij de drempel bepaald wordt door een kromme in de toestandsruimte. Met behulp van koppelingsargumenten kan de optimale strategie nog verder worden gespecificeerd.

Het volgende model wordt behandeld in Hoofdstuk 5. Een bediende ontvangt klanten, gecategoriseerd in klassen, ieder met hun eigen karakteristieken, zoals bedieningsduur, kosten en geduld. Om de monotonie van het systeem te bewaren gebruiken we de zogenaamde geleidelijke intensiteitsafknotting. De relevante structuureigenschappen zijn invariant voor de afkapping. Na deze te hebben aangetoond kunnen we hieruit concluderen dat een generalisatie van de $c\mu$-regel optimaal is.

Een praktische manier om eigenschappen aantonen via waarde iteratie is door middel van het zogenaamde gebeurtenis-gebaseerd dynamisch programmeren. Hierbij worden de veschillende gebeurtenissen apart gemodelleerd door operatoren, waarvoor propagatie-resultaten gelden. Bij afgekapte modellen ontstaan er gebeurtenissen met speciale afgeknotte intensiteiten. Van deze gebeurtenissen en de bijbehorende operatoren wordt in Hoofdstuk 7 een overzicht gegeven, inclusief een opsomming van de belangrijkste nieuwe propagatie-resultaten.

# Acknowledgements

This thesis would not have been established, but for the help of several people. Therefore, I would like to take this chance to thank them for their contributions.

The most important person has been my supervisor Floske Spieksma. With her warm personality she has given me faith and encouragement during my time as a PhD student. I have learned a lot from conducting research together. She has been an inspiration and has given me the right amount of freedom as well. I enjoyed working with her.

I am thankful to the Mathematical Department in Leiden and especially to Frank den Hollander for providing me with the opportunity to do research in his group. Furthermore, I would like to thank Sandjai Bhulai. He was the one who made me enthusiastic for this field. Moreover, I appreciated the collaboration in the last years.

The nice environment in Leiden has certainly contributed to a good completion of my PhD. Especially Laurens Smit, Martin Göll, Dwi Ertiningsih and Jan-Pieter Dorsman have been a pleasant companionship during my time in Leiden.

The other most important person is my wife Aryéle. I would like to thank her for her continuous love and support. She and our children form a great place to go home to after a days work. Finally, I would like to thank my parents for their everlasting support.

# Curriculum Vitae

Herman Blok was born in Utrecht, on June 22th, 1985. He completed high school at the Koningin Wilhelmina College, Culemborg, in 2003. He obtained his Bachelor's Degree in Mathematics at the Universiteit van Utrecht in 2010. In 2011 he obtained a Master's degree in Stochachtics and Financial Mathematics in a joint Master programme of the Universiteit van Amsterdam, the Universiteit Utrecht and the Vrije Universiteit van Amsterdam. His Master's thesis with the title 'Markov decision processes with unbounded transition rates' was written under the supervision of dr. Sandjai Bhulai and dr. Floske Spieksma.

He started his PhD at the Mathematisch Instituut van de Universiteit Leiden in 2011 under the supervision of Floske Spieksma. The result is presented in this thesis. His next position will be at the Technische Universiteit van Eindhoven with prof. Geert-Jan van Houtum and dr. Joachim Arts.